A Framework for Scheduling and Optimization of Healthcare Resources

A Thesis

submitted in fulfillment of the requirement
for the award of degree of

Doctor of Philosophy

Submitted By

Hemant Petwal
(Roll No.: 901503031)

Under the guidance of

Dr. Rinkle Rani
(Associate Professor, Computer Science and Engineering Department)
Thapar Institute of Engineering & Technology, Patiala

COMPUTER SCIENCE AND ENGINEERING DEPARTMENT
THAPAR INSTITUTE OF ENGINEERING & TECHNOLOGY
PATIALA - 147004, Punjab

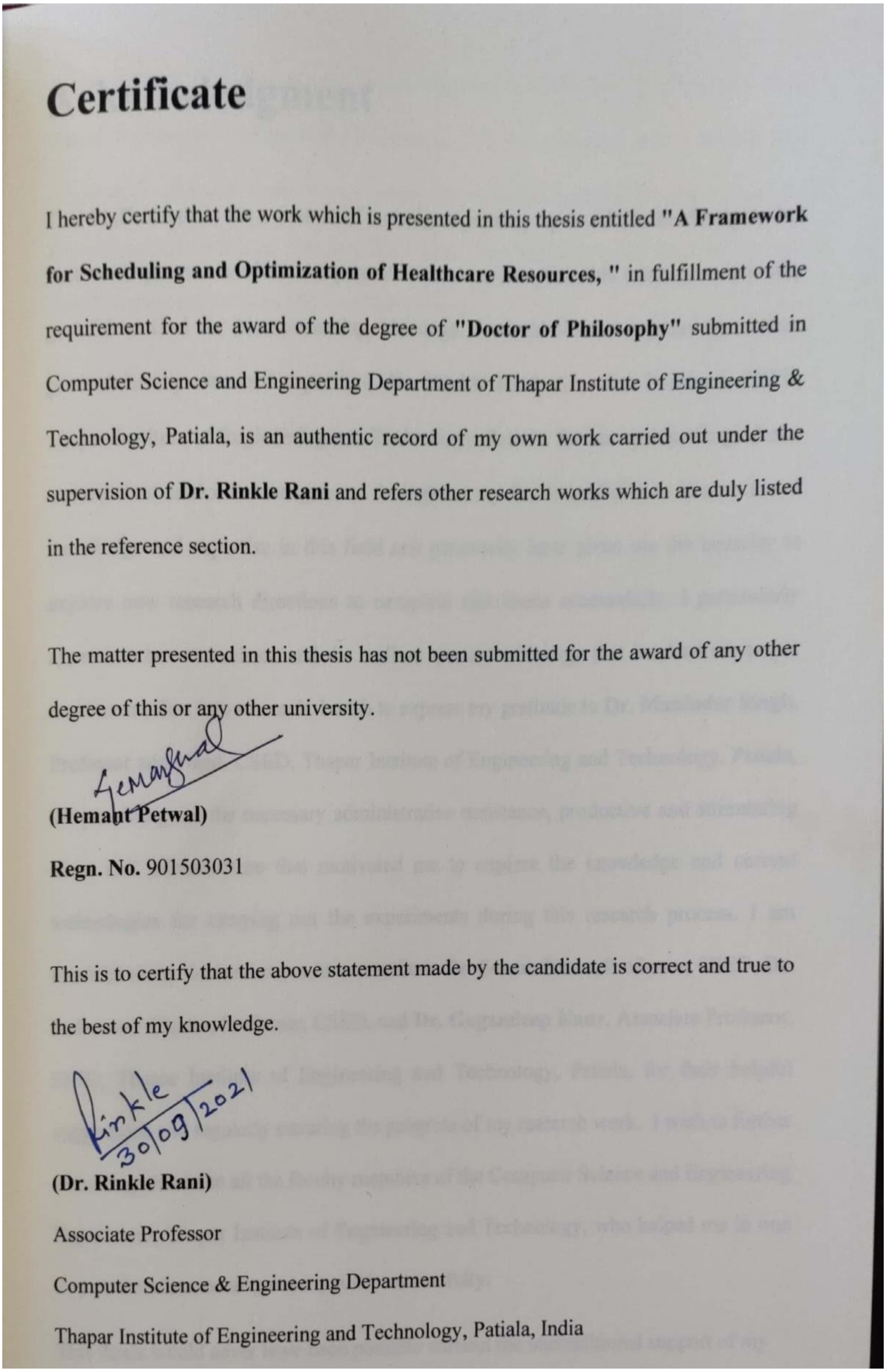
APRIL 2021

Contents

List of Figures

List of Tables

List of Algorithms



Certificate

I hereby certify that the work which is presented in this thesis entitled **"A Framework for Scheduling and Optimization of Healthcare Resources, "** in fulfillment of the requirement for the award of the degree of **"Doctor of Philosophy"** submitted in Computer Science and Engineering Department of Thapar Institute of Engineering & Technology, Patiala, is an authentic record of my own work carried out under the supervision of **Dr. Rinkle Rani** and refers other research works which are duly listed in the reference section.

The matter presented in this thesis has not been submitted for the award of any other degree of this or any other university.

(Hemant Petwal)

Regn. No. 901503031

This is to certify that the above statement made by the candidate is correct and true to the best of my knowledge.

Rinkle 30/09/2021

(Dr. Rinkle Rani)

Associate Professor

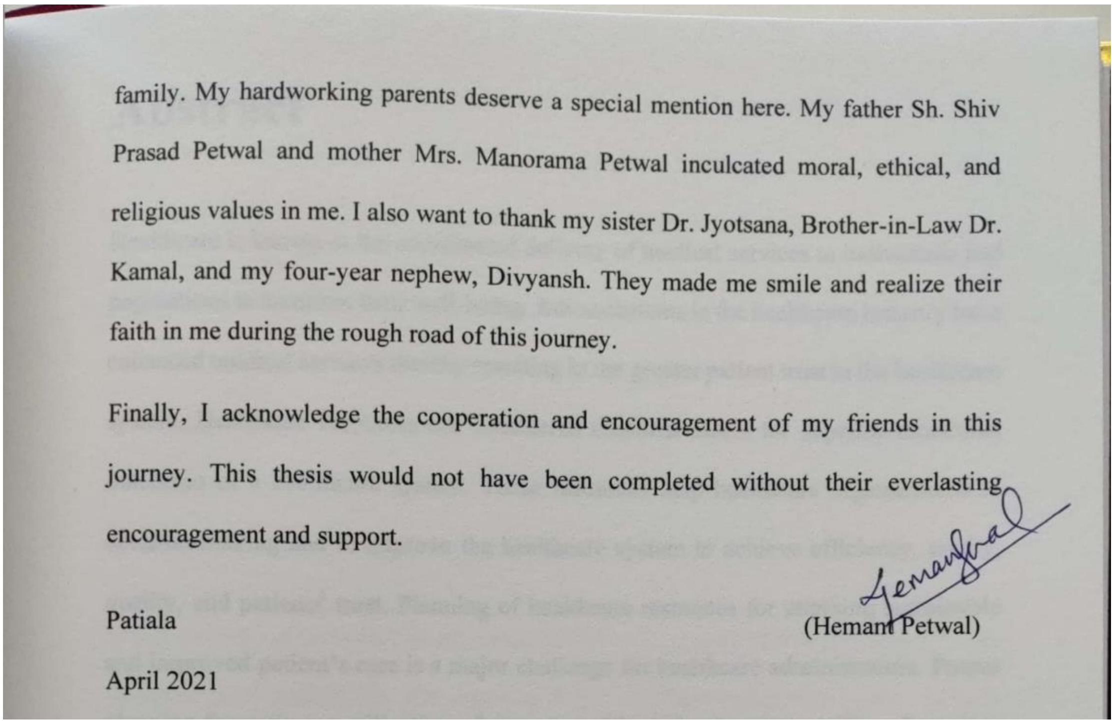
Computer Science & Engineering Department

Thapar Institute of Engineering and Technology, Patiala, India

Acknowledgment

This Ph.D. thesis is the outcome of an exigent journey upon which the Almighty has bestowed upon me wisdom and perseverance to accomplish this work. I am eternally grateful to God for his everlasting blessings. I would like to express my sincere gratitude to my esteemed supervisor, **Dr. Rinkle Rani**, Associate Professor, CSED, Thapar Institute of Engineering and Technology, Patiala, for her systematic guidance and constant advice encouragement throughout my Ph.D. candidature. Her profound knowledge and expertise in this field and generosity have given me the audacity to explore new research directions to complete this thesis successfully. I particularly appreciate her enthusiasm and the endless amount of energy that inspired me to put maximum effort into research. I wish to express my gratitude to **Dr. Maninder Singh**, Professor and Head, CSED, Thapar Institute of Engineering and Technology, Patiala, for providing me the necessary administrative assistance, productive and stimulating ideas and infrastructure that motivated me to explore the knowledge and current technologies for carrying out the experiments during this research process. I am thankful to my doctoral committee members: **Dr. Seema Bawa**, Professor, CSED, **Dr**. **Inderveer Chana**, Professor, CSED, and **Dr. Gagandeep Kaur**, Associate Professor, EIED, Thapar Institute of Engineering and Technology, Patiala, for their helpful suggestions and regularly ensuring the progress of my research work. I wish to further extend my thanks to all the faculty members of the Computer Science and Engineering Department, Thapar Institute of Engineering and Technology, who helped me in one way or the other to carry out my work successfully.

This thesis would never have been possible without the unconditional support of my



family. My hardworking parents deserve a special mention here. My father Sh. Shiv Prasad Petwal and mother Mrs. Manorama Petwal inculcated moral, ethical, and religious values in me. I also want to thank my sister Dr. Jyotsana, Brother-in-Law Dr. Kamal, and my four-year nephew, Divyansh. They made me smile and realize their faith in me during the rough road of this journey.

Finally, I acknowledge the cooperation and encouragement of my friends in this journey. This thesis would not have been completed without their everlasting encouragement and support.

Patiala (Hemant Petwal)

April 2021

Abstract

Healthcare is known as the coordinated delivery of medical services to individuals and populations to ascertain their well-being. Healthcare resources are considered essential assets for superior functional outcomes of a healthcare system. These resources help healthcare organizations in decision-making and to improve the healthcare system to achieve efficiency, equity, quality, and patients' trust. Planning of healthcare resources for attaining sustainable and improved patient’s care is a major challenge for healthcare administrators. Proper planning for optimum utilization of resources ultimately promotes quality of services and reflects the healthcare provider's performance and commitment towards patient care. Growing population is also contributing in the increased demand of timely medical services. Since healthcare resources are inherently limited, varying health needs generate disparities between increasing demand and the healthcare delivery process, thereby affecting the utilization and efficiency of health systems' resources and sustainability. Ultimately both health outcomes quality and healthcare performance suffer a lot. Hence, there is a requirement of methods and systems those can help in managing and utilizing healthcare resources to enhance patient care and overall performance. Scheduling and optimization of healthcare resources is an active research area now a days. Researchers working in the field of Scheduling and optimization of healthcare resources study the needs of patients and community, thereby explore how decisions should be made so that the demands of medical services can be aligned. Analysis of healthcare data is carried out to gain insights and making decisions for the overall management of resources. Researchers have applied various soft computing and conventional computation methods for

resource-related decision-making. In such data-driven analysis, evolutionary computing methods add a firm hand.

In this thesis, we have focused on scheduling and optimization of healthcare resources such as waiting list management, classification of patients and selection of healthcare personnel using soft computing approaches. The overall objective of this research is to develop a framework containing modules for surgical waiting list prioritization, surgical patients clustering, and surgical team selection.

In this research, a decision-making model PSWL-CCI is proposed to prioritize patients on the surgical waiting list. The proposed model addresses two critical issues: First, to prioritize patients from the surgical waiting list. Second, to refine and optimize the cosine consistency index (CCI) of inconsistent pair-wise comparison matrix (PCM) and obtain consistent priorities. The cosine maximization method (CM) with the analytic hierarchy process (AHP) is used to compute the priority of patients from the surgical waiting list, and a hybrid algorithm, HMWCA (Hybrid modified water cycle algorithm), is proposed to improve and optimize the cosine consistency index (CCI) of inconsistent pair-wise comparison matrix (PCM). The proposed hybrid algorithm exploits the features of three traditional algorithms, namely the evaporation-based water cycle algorithm (ER-WCA), genetic algorithm, and 2-opt heuristic algorithm. In the proposed algorithm (HMWCA), the concept of salt concentration and absorption is applied to the evaporation rate of ER-WCA that improves the modified water cycle algorithm (MWCA). The performance of the proposed algorithm is tested on different inconsistent PCMs and compared with existing algorithms. The optimized CCI values obtained by the proposed algorithm are validated through paired sample t-test also. Finally, the proposed model is validated through a case study of a real patient dataset

from an orthopedic surgery department of a multispecialty hospital in India. The proposed model is compared with existing prioritization methods. The experimental results reveal that the proposed model and associated algorithm significantly improve the CCI values and generate optimum priorities for the patients of the surgical waiting list.

Next, focusing on hospital’s existing surgical record management procedure (SRMP), we have proposed an efficient clustering algorithm to arrange patients in the optimal number of distinct groups on the basis of similarity of characteristics. The proposed clustering algorithm is based on population-based metaheuristic artificial electric field algorithm (AEFA) and is designed to deal with the mixed dataset. The proposed algorithm utilizes real-encoded variable-length cluster representation scheme to illustrate the candidate solution, which enables the algorithm to find optimal number of clusters automatically. The concepts of threshold setting and cut-off ratio are used in the optimization process to further refine the clusters. In the proposed algorithm, similarity among data points and different cluster centers is measured using Euclidean distance (for numeric attributes) and probability of co-occurrence (for categorical attributes). The performance of the proposed algorithm is tested using real-life datasets and compared with existing mixed data clustering algorithms on the basis of two parameters: average accuracy and standard deviation. The proposed algorithm is validated using an unpaired t-test also. Finally, the proposed algorithm is validated using a case study of real postoperative surgical mixed data set obtained from the surgical department of a multispecialty hospital in India. It is observed from the results that the proposed clustering algorithm arrange the patients into optimal subgroups efficiently.

Finally, a decision-making model is proposed to select an optimal list of surgical teams. The proposed model addresses two critical issues: First, to improve the existing surgical history management system of a multi-specialty hospital (MSH), and second, to select an optimal list of surgical teams for a new surgical patient. Therefore, two modules: the surgical history management (SHM) module and the surgical team selection (STS) module, are introduced in the proposed model. The SHM module aims to arrange surgical patients into optimal sub-groups on the basis of similar characteristics. It helps the STS module in the selection of the optimal list of the surgical team for a newly referred surgical patient. In the STS module, a population-based meta-heuristic algorithm (AEFA) for multi-objective optimization is proposed to select the optimal surgical team. The performance of the proposed clustering algorithm is tested on real-life datasets and compared with the existing clustering algorithm. Further, the performance of proposed multi-optimization algorithm tested on benchmark functions and compared with existing multi-objective optimization algorithms. Finally, the proposed model is applied to a real postoperative surgical dataset from the orthopedic surgery department of MSH in India. The experimental results show that the proposed model proves its efficiency in selecting the optimum surgical teams for a patient.

Keywords: Multi-Criteria Decision-Making, Pair-wise Comparison Matrix, Cosine Consistency Index, Water Cycle Algorithm, Genetic Algorithm, 2-Opt Heuristic, Priority Vector, Data clustering, Meta-heuristic, Artificial electric field algorithm, Distance measure, Mixed dataset clustering, Multi-objective optimization, Surgical team, strength Pareto.

Chapter 1

INTRODUCTION

This chapter presents introduction of a healthcare system, its components and functional structure of a healthcare organization. Subsequently, the classification of various healthcare resources and the importance of resource optimization are discussed in detail.

Health is defined as a state of complete physical and mental well-being free from any disease or debility[1]. Healthcare is known as the coordinated delivery of medical services to individuals and populations to determine and improve their well-being. It includes the maintenance and improvement of patient mental and physical health through the diagnosis, treatment, and prevention of illness, injury, and other physical and mental impairments. Healthcare pertains to a system known as the healthcare system. The Healthcare system involves health care providers and resources that provide health care services to meet the target population's medical needs. The health care providers are the organization or healthcare professionals responsible for delivering healthcare services such as nursing care, outpatient department, emergency care, surgical care, pharmacy and diagnosis, ward facility, etc. to the patients. An authority administers a healthcare system is called health care management that provides leadership and direction to providers to deliver personalized health services. Advancements in the healthcare system have improved healthcare delivery as per patients’ health needs, contributing to greater patient trust in the healthcare system. In a growing population, hospitals’ primary concern is to provide quality healthcare through the effective use of healthcare resources to satisfy patients[2]. To deliver quality services, hospitals implement various strategies ranging from efficient use of existing infrastructure and resources to upgrading infrastructure, resources, and services[3]. An overview of healthcare system is presented in Figure 1.1. Hospitals focus on care, cost, and delivery and implement different strategies for patients to deliver quality services, ranging from utilizing existing infrastructure and resources efficiently to upgrade the infrastructure, resources, and services. Optimal resource utilization and well-scheduled care delivery are seen as the future of the healthcare system[4]. It helps healthcare organizations (providers) in enhancing the distribution and utilization of healthcare resources to increase the service capacity, boost hospital performance, and minimize access to healthcare services for patients. Researchers working in the field of healthcare optimization and scheduling study the needs of patients or the community and explore how decisions should be made so that the demands of the population can be aligned[4]. They are also studying hospital functional aspects and identifies the factors influencing performance. The factors[5] influencing hospital performance and patient satisfaction lay the foundation of various kinds of healthcare data analysis that can be conducted using statistics and computational approaches. Healthcare data analysis is widely utilized to gain insights about patients as well as the hospital to assist in healthcare decision-making such as prioritized patient care, patients’ clustering, healthcare personnel selection, etc. Researchers and scientists contribute to exploring the different approaches that assist the healthcare system by making the optimal use of resources and scheduling them effectively. In such data-driven analysis, evolutionary computing methods add a firm hand.

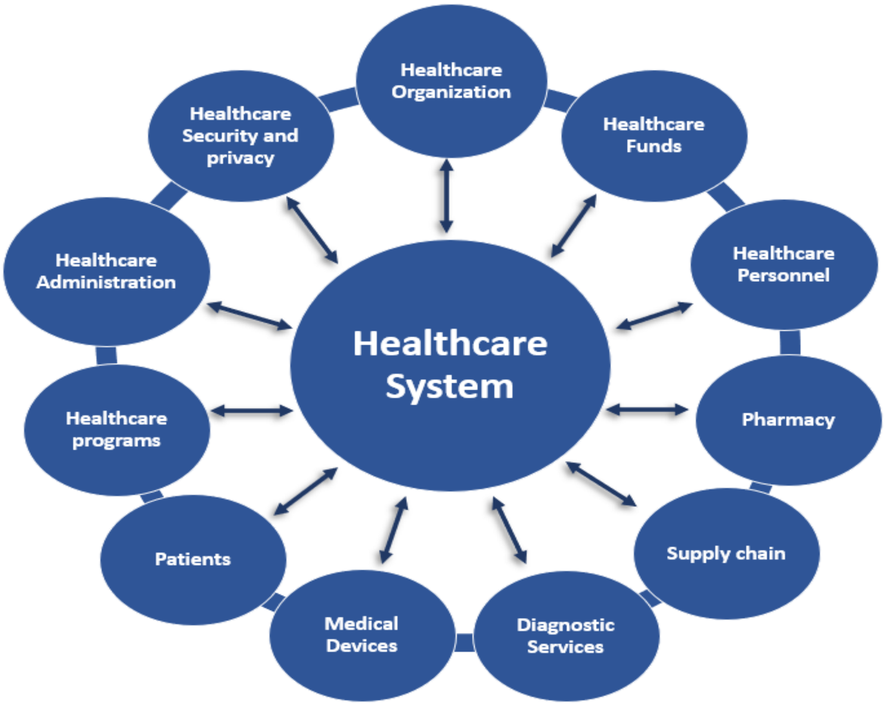


Figure 1.1: Components of healthcare system

1.1 Functional Structure of Healthcare Organization

Every healthcare organization, also known as healthcare providers, has a functional structure to treat their patients. The structure (Figure 1.2) is discussed as follows:

1.1.1 Patients’ Check-In

Patients’ check-in is the process where patients register themselves with the healthcare facility to receive health-related services. Patients are classified into the following type:

A. Inpatient

A patient who has to stay in the hospital for one or more nights or several days or weeks or in certain severe conditions such as coma and day-to-day check-ups for many years or until death is defined as an inpatient.

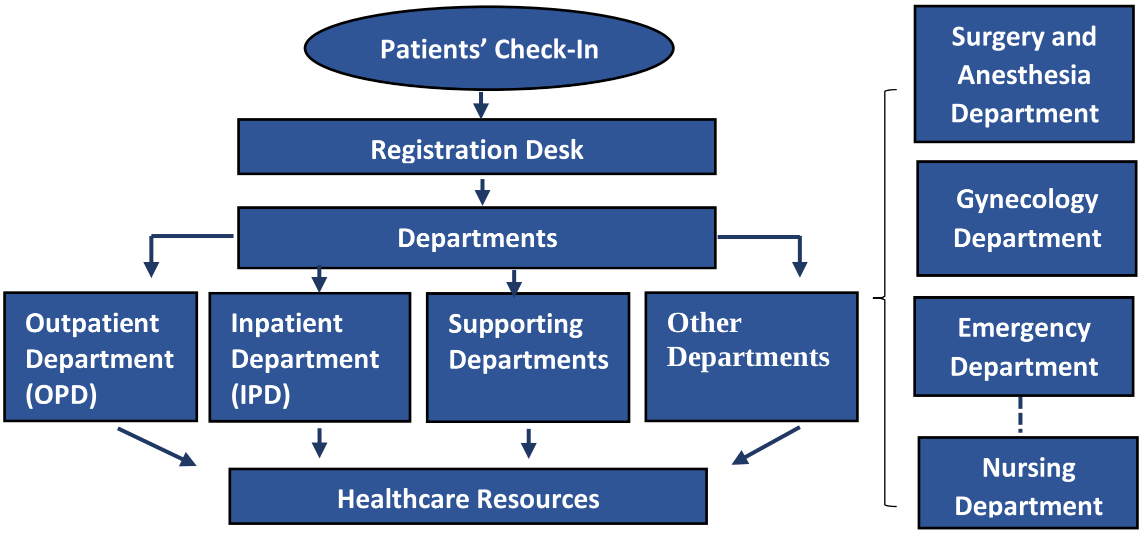


Figure 1.2: Functional structure of healthcare organization

B. Outpatient

An outpatient is a patient who visits the hospital without an intention to stay longer than the length of the appointment.

C. Day Patient

A day patient is a patient who remains throughout the day in a hospital or clinic to access medical services, such as minor surgery or other therapies. These patients do not expect to stay in the hospital overnight.

1.1.2 Registration Desk

This level handles registration, admission of patients, billing, medical record, computer-generated information, etc.

1.1.3 Departments

After registration, patients are referred to their respective departments for medical consultation or treatment. A hospital contains various departments, according to its size and capacity, are as follows:

(i) Outpatient Department (OPD)

This department treats all patients who visit the hospital, clinic, or associated facility for diagnosis or treatment but do not require to stay there for overnight care.

(ii) Inpatient Department (IPD)

This department offers care to patients who stay overnight or longer for medical treatment in the hospital.

(iii) Supporting Departments

It involves departments that provide supportive services, e.g., catering and food department, laundry service, laboratory/pathology department, blood bank department to inpatients.

(iv) Other Departments

It involves the surgery & anesthesia department, gynecology department, emergency department, nursing department, etc.

1.2 Healthcare Resources

Resources are the key component of a healthcare system. All personnel, materials, equipment, facilities, funds, and knowledge used to provide health care services are called healthcare resources[6]. A healthcare system is always committed to delivering quality health services to patients and relies on the availability of healthcare resources to operate these services. Healthcare resources facilitate the running of the entire system systematically and effectively overall. The more the healthcare system is equipped with high-quality resources, the more it can provide its patients with outstanding care. The healthcare resources are classified as follows (Figure 1.3):

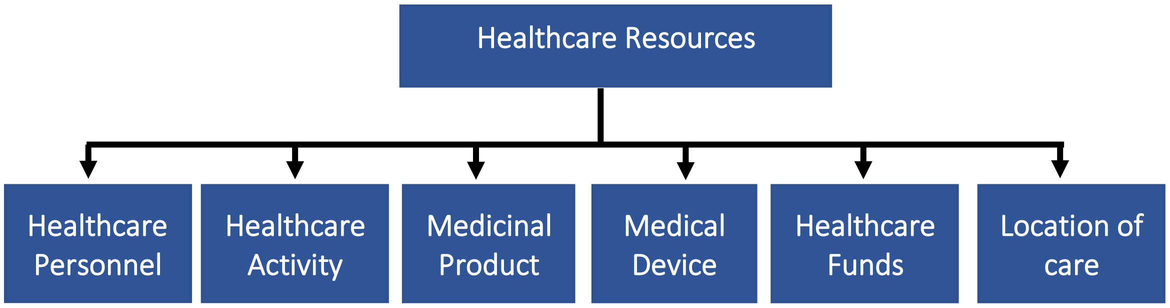


Figure 1.3: Classification of healthcare resources

1.2.1 Healthcare Personnel

Health care personnel[7] are persons with special education in health care who are closely involved with providing health care services. Physicians, surgeons, nurses, nursing assistants, therapists, technicians, emergency medical service personnel, dental personnel, pharmacists, laboratory personnel, autopsy personnel, students, trainees, etc., involved in the healthcare system fall under this category. Whether or not directly interact with the patients, these healthcare workers are involved in dispensing their health-related medical advice, treatment, and diagnosis to patients. Further, individuals who are not involved in patient care directly, such as clerical, housekeeping, dietary, laundry, maintenance, billing, and volunteers, but employed in the healthcare system are also known as healthcare personnel. Such personnel interact with patients but are not involved in assisting any medical aspects.

1.2.2 Health Activity

Healthcare activity is defined as the process intended to improve or maintain the health status of patients. Healthcare activities[8, 9] are classified as: 1) direct care and 2) indirect care. Direct care involves activities in which healthcare personnel spends time with the patient, e.g., consultation and counseling of patients, patient education, diagnosis, etc. Indirect care involves activities performed on behalf of patients but doesn’t require interaction with patients, e.g., Scheduling an activity, documentation, post-procedural report generation, etc.

1.2.3 Medicinal Product

Medicinal products are defined as substances or combinations of substances intended for the treatment, prevention, and diagnosis of illness to cure physiological functions by pharmacological, immunological, or metabolic action.

1.2.4 Medicinal Devices

Medical devices are defined as any software, instrument, apparatus, machine, appliances, implant, etc. intended to be used, alone or in combination, for human beings for one or more specific purpose(s).

1.2.5 Healthcare Funds

In healthcare, the health fund is an account associated with any health insurance plan for the patients that helps them pay for received healthcare services.

1.2.6 Location of Care

Location of care is defined as a place where patients are delivered healthcare services.

It involves operation theaters, diagnostic centers, ambulatory surgical centers, intensive care units, etc.

Healthcare resources are considered to be crucial assets for various functional outcomes of a healthcare system. It helps healthcare organizations in decision-making to reform the healthcare system to achieve efficiency, equity, and quality. Efficiency implies avoiding the overuse or misuse of the resource in the healthcare delivery process. Equity is defined as the fair distribution of care process to the target population so that everybody can have a reasonable chance to receive health services. Quality means delivering care to patients safely and effectively when needed to improve health and gain patients' satisfaction. Various computation methods, machine learning[10] methods, soft computing techniques, and analytic[11] tools use healthcare resource data to gain decisive insights to enhance healthcare organizations' performance. Active research on health care resources has exploited its potential in exploring new research possibilities to reform the healthcare system[12,13].

The aging population and varying health needs affect the sustainability of health systems. Since health care resources are inherently limited, the utilization and efficiency of resources are affected, resulting in disparities between increasing demand and the healthcare delivery process[14, 15]. Ultimately both health outcomes quality and healthcare performance suffer a lot. Hence, it is required to develop methods or systems that can help manage and utilize healthcare resources to enhance patient care and healthcare performance. Different approaches have emerged in the last decade in the direction of efficient management and utilization of healthcare resources[16, 17]. Researchers are using various soft computing and conventional computation methods for resource-related decision-making[18-20]. Researchers are also developing computation models/algorithms[4] that can help efficiently manage healthcare resources by prioritizing patient access, clustering of patients, and selecting healthcare workers.

1.3 Role and Application of Optimization in Health care

Decision-making on resources and care delivery processes to sustain or increase care quality is one of the fundamental challenges for healthcare operations management. The right decisions promote the importance of service quality and reflect the healthcare provider's commitment to patient care. Contrary, wrong decisions lead to irreparable results on the quality of services and adversely affect healthcare providers' performance. Optimization is among the several widely utilized operational research management techniques to address such complex healthcare management problems[4]. Optimization is among the several widely utilized operational research management techniques to address such complex healthcare management problems. Optimization assists in decision-making by determining the best possible way to focus on the healthcare delivery process. Scheduling and optimization help prepare, manage, and utilize healthcare resources efficiently to ensure that the different healthcare delivery phases align with the patient’s needs. Both Scheduling and optimization have been widely used in several healthcare fields, ranging from planning the organizational level of healthcare decisions to developing national healthcare policies.

1.3.1 Waiting-List Management

A waiting list is a queue of patients awaiting treatment. Waiting list management is a part of appointment scheduling where patients in the queue are scheduled for treatment in an appropriate order based on the patients' condition and healthcare availability so that resources can effectively be used and patients' waiting times can be reduced. Since the uncertainty in health care service time is a vital challenge in the healthcare delivery process, optimization is utilized in waiting list management. Optimal waiting list management involves minimizing waiting time, prioritizing patients, optimal utilization of resources, etc. Considering surgical patients especially, the waiting time for a patient to undergo surgery is a critical problem. When a patient is referred for surgery, he is first examined and then admitted to an appropriate department. This type of admission is known as "first come, first treatment," where patients receive treatment based on the time or sequence of their admission. However, patients with severe health conditions get accommodated on a priority waiting list, which does not follow the "first come, first treatment" basis. In simple words, each patient receives treatment on a priority basis. Patients with high priority receive immediate medical attention in comparison to patients with low priority. Prioritization of patients, according to their medical condition, is a complex decision-making process in hospitals and medical institutions. Assessing patients' conditions on the waiting list and providing them with the appropriate medical facilities on a priority basis are two critical issues. The prolonged waiting period often deprives patients of timely medical attention, which further exacerbates their condition. Besides, it is a traumatic experience for the friends and family members of the patients as they witness the patient fight the agony of the prolonged waiting period. Prolonged waiting time detriments the health of a patient. Over the past few years, the waiting time has emerged as a critical issue in north-western and north-eastern countries. Patients on waiting lists could not be treated immediately for various reasons. It includes the lack of medical resources (surgeons, doctors, and nurses), long waiting queues, and poor management of waiting lists.

Subjective errors and inappropriate assessment of the patients' medical condition were two reasons for the poor management of the waiting list. Existing prioritization methods such as a priority scoring system[21], homogeneous waiting group approach[22], and analytic network process[23] guided doctors on how to perform priority-based decision-making. Although existing priority assessment systems have proven useful, these systems display irregularities in the evaluation of priorities. Some prioritization systems show uncertainty in the judgment of decision-makers[24]. Some methods help health care providers in better decision-making[25]. Moreover, some of these systems neither support group decision-making[26] nor provide relationships between criteria[27] related to priority evaluation. A brief review of the shortcomings of the currently existing systems is provided in the literature[28].

1.3.2 Surgical Patient Clustering

Surgical patient clustering is defined as a process of arranging the patients into distinct groups based on their similarity of characteristics. These characteristics include age, gender, body mass index (BMI), American Society of Anesthesiologists (ASA) fitness grade, etc. For multi-specialty hospitals, where an enormous number of patients receive their surgical care, it is quite challenging to manage the surgical records efficiently. Efficiently managed surgical records help hospitals to improve patients care and monitoring to enhance the efficiency of resources within the hospital.

Data clustering is a data analysis approach that arranges unlabeled data into different groups based on a similarity measure. Each group is called a "cluster," which shows similarity among the data in it and differs from the set of data in other clusters. Clustering is most widely used in those disciplines where multivariate data analysis is required. In recent years, cluster analysis has played a significant role in the distinct domains of various fields such as engineering, life, and medical sciences, earth sciences, and economics[29, 30, 31]. The elementary problem of the clustering analysis is to accurately determine the approximate number of clusters, as this number influences the clustering outcomes to a large extent[32]. Clustering algorithms are classified into two main categories: partitional clustering[33] and hierarchical clustering[34]. The hierarchical clustering arranges data points in a hierarchical tree structure based on the homogeneity among the data points. It overlooks the shape and size of the formed clusters. Further, this clustering allocates a single cluster to a data point that renders the cluster structure static. Contrary, partitioning clustering analyzes the dataset and organizes data points into clusters based on the similarity among data points. The partitioning clustering aims to optimize a global criterion involving minimizing the similarity among the elements within a cluster and maximizing the disparity between different clusters. Although both of these algorithms prove their usefulness and performance in various domains, both still have some critical limitations. The efficacy of these algorithms depends on the foreknowledge of the number of clusters present in the datasets. Since different datasets, especially in real-world applications, have diverse patterns, cluster analysts lack information on how many appropriate clusters exist in the dataset[35]. Hence, these algorithms that require the cluster number as an onset parameter cannot be used effectively. Since most real-world datasets do not have class labels, there are no specific criteria for directing clustering analysis. It is considered a major limitation[32] of the dataset, which makes it a challenging task to find a suitable number of clusters. Therefore, determining the optimum number of clusters in a data set has become an essential research issue to address such limitations. In recent years, the concept of the automatic clustering method has been used to overcome this limitation in clustering. Automatic clustering is defined as an analytic process of determining a suitable number of clusters in the dataset, irrespective of any prior knowledge related to the dataset[36]. Several automatic clustering algorithms, mostly inspired by natural phenomena, involving genetic algorithm (GA)[37, 38], particle swarm optimization (PSO)[39, 40], gravitational search algorithm (GSA)[41, 42], differential evolution (DE)[43], bacterial evolutionary algorithm (BEA)[44], and bee colony optimization algorithm (BCA)[45, 46], etc., are introduced in recent years. In these algorithms, clustering is considered an optimization activity, aiming to maximize the similarity among a cluster's data and maximize the disparity among disjoint clusters[47]. These algorithms have proven their higher convergence speed and efficacy in producing quality outcomes not only in optimization but also in clustering analysis.

1.3.3 Capacity Planning

In health care, capacity planning involves decisions related to three types of resources: healthcare workforce, diagnostic or surgical equipment, and facilities such as beds, ambulance, etc. Focusing on the healthcare workforce, capacity planning is an important process in surgical decision-making. To meet patients' demands, enhance surgical outcome quality and patient satisfaction, optimization assists capacity planning to optimize and acquire these resources.

Considering the healthcare workforce, during a preoperative procedure, surgical team members, surgical specialty and experience, and coordination among the members play essential roles[48]. Although various factors[48] can affect surgical procedures, positive outcomes mainly depend on the individual surgical team members. Appropriate coordination and cooperation among those can reduce unavoidable conflicts during the procedures[49]. Hence, the selection of an optimal surgical team is indispensable for a rapid patient recovery, decreased complications, and more favorable surgical management. However, the selection process is a considerably time-consuming and difficult task[50]. In recent years, several studies have examined the performance of surgical team members. Many vital factors, such as understanding diagnostic complications and patient characteristics, and the surgical practice environment, can result in more satisfactory outcomes. In the field of medicine, appropriate management of surgical care is a difficult task[51, 52] as most surgical complications occur during intra operative surgical care[53]. An efficient team can help in providing effective healthcare services[54]. Hospitals and physicians always focus on providing a safe environment for patients and enhancing their well-being[55]. A study investigated human factors associated with operating rooms and analyzed the relationship between their poor performance and the surgical procedures outcomes[54, 56]. Similarly, several studies adopted approaches such as malpractice claim analysis[57, 58], root cause analysis[59, 60], and prospective analysis[61, 62] to reduce intraoperative surgical complications. Although these studies tried to analyze the relationship between the performance of operating room and outcomes of surgery, however, contribution of significant factors affecting the performance of operating room were not considered[53]. Studies examining factors such as teamwork in the operating room[63, 64], trauma care[65], and intensive care[66] focused on the effect of coordination and synergy among surgical team members. These studies have indicated the necessity and significance of surgical team selection procedure. As performing surgical procedures is often a risky and uncertain task, therefore high synergy is always expected among team members possessing different levels of experience and expertise. Various factors such as availability of surgeons, limitation of resources, and time etc. affect the surgical team selection in a multispecialty hospital; thus, selection of the surgical team is a challenging task[67]. In a surgical team, different responsibilities are assigned to different individuals[68]. As the responsibilities and the individuals to whom responsibilities are assigned change frequently with time, thus selecting an efficient team for the desired activity, considering time and resource limitation, becomes a complicated procedure[68]. All the aforementioned studies have focused on personnel preferences for the day, shift, and units. However, none of these studies have considered the history of the surgical team, characteristics of patients, and feedback of patients who underwent surgery in the past.

1.4 Thesis Organization

The thesis is organized as per the following chapters:

Chapter 1: Introduction

This chapter introduces healthcare, its functional structure, and its resources. Subsequently, the classification of healthcare resources and the role of optimization in healthcare are discussed in detail. Further, some applications of optimization in managing healthcare resources are discussed in this chapter.

Chapter 2: Literature Survey

This chapter presents a detailed review of the literature related to approaches available for scheduling and optimization of healthcare resources.

Chapter 3: Problem Formulation

This chapter discusses the research gaps, defined research objectives, and methodology adopted. The research gaps are identified from the existing literature related to patient prioritization, patients clustering, and surgical team selection. By considering these gaps, four research objectives are identified for healthcare resource optimization and scheduling. The third objective is further divided into three sub-objectives: patient prioritization from the surgical waiting list, postoperative surgical patients clustering, and optimal surgical team selection. Additionally, the methodology used for achieving various objectives is given in detail.

Chapter 4: Prioritizing Patients from the Surgical Waiting List

This chapter discusses challenges related to patient prioritization usually faced by surgeons while attending patients who are waiting for surgery. It explains the proposed decision-making model to prioritize patients of the surgical waiting list. The proposed model addresses two critical issues: First, to prioritize patients from the surgical waiting list. Second, to refine and optimize the cosine consistency index (CCI) of inconsistent pair-wise comparison matrix (PCM) and obtain consistent priorities. The perception of surgeons for identified parameters in the term of rating is considered to determine priorities from the given surgical waiting list. The cosine maximization method (CM), along with the analytic hierarchy process (AHP), is used to compute the priority of patients from the surgical waiting list. To improve the inconsistent pair-wise comparison matrix (PCM), a novel hybrid algorithm, HMWCA (Hybrid modified water cycle algorithm), is proposed and incorporated in the model. The proposed algorithm exploits feature of three traditional algorithms, namely evaporation-based water cycle algorithm (ER-WCA), genetic algorithm, and 2-opt heuristic algorithm. The performance of the proposed algorithm is tested on different inconsistent

PCM and compared with existing algorithms. The optimized CCI values obtained from the proposed algorithm are validated through paired sample t-test also. Finally, the proposed model is validated through a case study of a real patient dataset from an orthopaedic surgery department of a multispecialty hospital in India. It is shown through the experimental results that the proposed model and associated algorithm significantly improve the CCI values, thus generate optimum priorities for the patients of the surgical waiting list.

Chapter 5: Cluster Analysis of Surgical Patients

This chapter focuses on finding the similarities among the surgical patients and discusses the proposed clustering algorithm for this purpose. The proposed algorithm is based on a population-based metaheuristic algorithm called artificial electric field algorithm (AEFA) and deals with the mixed datasets. This algorithm aims to cluster surgical patients into optimal partitions. The proposed algorithm utilizes (i) a real-coded variable-length candidate solution to detect the optimal number of clusters automatically, (ii) the concepts of threshold setting and cut-off ratio in the optimization process to refine the clusters, and (iii) Euclidean distance (for numeric attributes) and the probability of co-occurrence of values (for categorical attributes) for measuring the similarity between data points and different cluster centers. The proposed algorithm is compared with existing mixed dataset clustering algorithms and validated on publicly available real-life datasets, namely, heart disease, credit approval, iris, and soybean using two robustness measures: average accuracy and standard deviation. The proposed algorithm is applied to a real postoperative surgical mixed dataset obtained from a multispecialty hospital's surgical department.

Results show that the proposed algorithm outperforms the existing mixed dataset clustering algorithm and organizes the surgical patients into an optimal number of partitions.

Chapter 6: Optimal Surgical Team Selection

In this chapter, considering the challenge of the selection of a surgical team faced by multispecialty hospitals (MSHs) in the healthcare system, a decision-making model is proposed to select the optimal list of surgical teams for a patient. The proposed model deals with two critical issues; (i) to improve the existing surgical history management system by arranging surgical patients into optimal sub-groups, and (ii) based on optimal sub-groups, select an optimal list of surgical teams for a new surgical patient. Two population-based meta-heuristic algorithms for clustering of mixed dataset and multi-objective optimization are proposed for arranging surgical patients into optimal sub-groups and optimal surgical team selection. The proposed model is validated through a real postoperative surgical dataset from an MSH orthopedic surgery department in India. The experimental results show that the proposed algorithms prove their efficiency in selecting the optimum surgical teams for patients.

Chapter 7: Conclusion and Future Directions

This chapter concludes the thesis by providing a brief overview of the proposed framework/algorithms and provides insight into the future work scope.

Chapter 2

LITERATURE SURVEY

This chapter presents a detailed review of the literature related to computational approaches for prioritizing patients from the surgical waiting list, cluster analysis of surgical patients, and optimal surgical team selection. Section 2.1 presents a comparison of existing prioritization methods. Section 2.2 shows a comparative summary of existing clustering algorithms. Section 2.3 provides an overview of existing methods used in surgical decision-making.

2.1 Prioritizing Patients from the Surgical Waiting List

One of the most complex tasks for medical practitioners is to evaluate patients' priority for treatment[69]. Generally, patients whose conditions demand urgent intervention by doctors or who claim their position in the priority list were assigned priorities. Some of the researchers[22, 70, 71, 72] proposed that ascertaining the severity of medical conditions helps determine the patients' priority. A proper scoring system was used to help ascertain which patients need to be treated on a priority basis. A scoring system was a procedure in which weights were assigned to symptoms of a disease. Patients were then prioritized by associating them with the symptoms of certain diseases and determine the severity of their condition. A comprehensive study was conducted to assess which patients need to be given priority[73]. This study used a model that was designed for elective surgery and relied on two essential factors; clinical urgency and the waiting time of each patient. Moreover, researchers were giving importance to scoring systems and urgency-rated groups to evaluate patients on a priority basis[21]. Several studies evaluated patients' priority by using medical data of patients requiring organ transplants[23, 27, 74, 75, 76], coronary artery bypass grafting surgery[77], and trauma surgery[78], as well as patients in the intensive care unit[79]. Recently, a study was conducted on surgical patient prioritization using AHP and the analytic network process. For priority evaluation of patients, a pair-wise comparison matrix was implemented[22]. All these researches helped in identifying a good number of approaches to determine priority. These included the weighted least squares (WLS) method[80], Eigenvector (EV) method[81], Additive normalization (AN) method[82], Logarithmic least squares (LLS) method[83], Eigen weight (GE) method and Least distance (LD) approximation method[84], and Fuzzy preference programming (FPP) method[26]. The different prioritization methods produced varying results as each of them perform differently[82, 84]. For instance, for measuring consistency, prioritization methods such as LLS, WLS, FPP, and EV showed better performance than goal programming (GP). In the case of measuring conformity, WLS, GP, LLS performed better than FPP. Similarly, LLS showed minimum violation and performed better than the rest of the prioritization methods. Henceforth, it is not easy to identify the best method that can perform well in all scenarios (Table 2.1). It makes clear that determining the superiority of prioritization methods is a challenge. In a PCM, the consistency and reliability of derived priorities are of extreme importance. If the PCM is inconsistent, it may result in inconsistent priorities; however, several methods exist to solve inconsistent multiplicative preference[86, 87, 88, 89]. To bring improvements in the cosine consistency index in AHP, G. Khatwani, and A. K. Kar focused on the maximum deviated element in inconsistent PCM[90]. Similarly, ACO was combined with AHP to bring improvements in inconsistent PCM[91]. Two main aspects that were the focus of the study included consistency threshold and distance between modified PCM and original PCM. Further, the consistency was improved by S. Siraj et al. to detect and remove the intransitive in comparison matrices[87]. The inconsistent matrix was modified by Ergu et al. with the help of proposed induced matrix to identify the elements leading the matrix to be inconsistent[92]. Consequently, they proposed to change the elements to obtain a consistent matrix. However, they decided not to change most of the elements of matrix except the entries that were suspected of making the matrix inconsistent. In addition to this, the use of an intelligent algorithm to modify an inconsistent matrix was a common research trend. The genetic algorithm (GA)[93, 94] was used to obtain the consistent matrix. Recently GA was utilized, with deep learning[95] for parameter optimization[96] and weight optimization[97], for cancer classification and detection. To modify inconsistent matrices, particle swarm optimization (PSO) was combined with the Taguchi method[98]. Two important aspects were also deduced by Lin et al. [94] and Yang et al. [98]. This included the distance between modified and original matrix, also known as difference index (D_i), and the eigenvalues were close to number of comparison elements, $max = n$. In order to minimize the overall index (OI), these aspects were integrated with the objective function. It is worth mentioning here that no significant consideration was given in the methods discussed above to find the lowest different matrix between the original (inconsistent) and modified (consistent). Although GA[94] and PSO[98] methods reduced the difference index, they do not help in getting the optimal difference index. Further, the above-mentioned methods have proven their usefulness still these method lacks potential in modifying inconsistent matrices. Most of these methods suffer from high computational time, high deviation between original and modified matrix, and need maximum number of iterations to satisfy the objective. Therefore, focusing on the fact that the matrix is more important if it is closer to the original matrix and the consistency of the matrix is acceptable, there is a need to develop an effective algorithm to modify the incompatible matrix. In a recent study, the water cycle algorithm (WCA)[99] has emerged as an intelligent nature-inspired optimization algorithms. There has never been any research to modify the inconsistent pair-wise comparison matrix by using WCA. WCA is considered to be a reliable tool to solve different optimization problems, including traveling salesman problem (TSP)[100], clustering[101], and scheduling[102]. A comparative summary of the existing prioritization methods presented in Table 2.1.

2.2 Clustering Analysis of Surgical Patients

K-means is a center-based clustering approach widely used in the partitioning of the data set due to its simplicity and efficacy. The major drawback of the k-means algorithm is its dependency on an initial number of cluster centers and its faster convergence to the local optima[104, 105]. Although several clustering algorithms[106-108] have been proposed to prevent a solution from being stuck in the local optima, the reliance of the algorithm on prior cluster number information and its adverse effect on clustering performance has emerged as an issue[109-112]. For addressing this issue, several optimization algorithms have been adapted to implement automatic clustering analysis. The first-ever contribution toward automatic clustering was based on evolutionary computing, called EP-clustering[113]. EP-clustering aimed to minimize the DB and WGS indices to improve the efficiency of exploration and exploitation. EP-clustering algorithm produced better results as compared K-means algorithm when implemented on real-life datasets. In recent years, Chen et al. [114] introduced a GEP-cluster algorithm

Table 2.1: Summary of existing prioritization methods

Author (s)	Objective/Work Done	Technique proposed/Used	Performance parameters	Research gap(s) Identified
Testi et al.[21]	Determined the priority of each surgical patient from the waiting list and the relative order of admission based on urgency	Prioritization scoring algorithm	1. Efficiency of algorithm 2. Equity of algorithm	This study lacks validation of proposed urgency assessment
Mariotti et al.[22]	Improved the prioritization of outpatients waiting for a specialist or a diagnostic test	Homogeneous waiting time groups approach	Number of referrals	This study focused on outpatient's referral only and not considered inpatients
Rahimi and Jamshidi[23]	Prioritized organ transplants surgical patients on the waiting list	Analytic network process	Robustness	Uncertainties were found between the judgments of decision-makers and converted crisp values
Rahimi et al.[24]	Prioritized elective surgical patients considering risk and uncertainties	1. AHP 2. Data envelopment analysis	Sensitivity analysis	Vagueness and inherent uncertainties found with the judgment of decision-makers
Mikhailov[26]	Evaluated the priority vector from a pair-wise comparison matrix (PCM)	Fuzzy Programming method	Euclidean distance	This study lacks an interval comparison between fuzzy programming and group decision-making

Lin and Harris[27]	Designed a framework to prioritize patients for organ allocation	Analytic hierarchy Process	Sensitivity analysis	This study lack interaction among different factors (efficiency, equity, urgency)
Lack et al.[71]	Proposed a point scheme to prioritize elective patients from the waiting list	1. Priority scoring system 2. Point scheme	1. Waiting time 2. Efficiency	The proposed system was more susceptible to favoritism in priority decision making
Al-Ebbini et al.[75]	Determined which potential recipients would receive a lung for transplantation as it becomes available	The fuzzy lung allocation system	Accuracy	It only mimicked real cases observed at UNOS (United Network for Organ Sharing) and was not intelligently adapted to new cases
Khatwani and Kar[90]	Improved cosine consistency index for the analytic hierarchy process for solving multi-criteria decision-making problems	1. Weighted arithmetic mean (WAM) 2. Weighted geometric mean (WGM)	Consistency index	This study not considered the deviation between the modified PCM and the original PCM
Girsang et al.[91]	Proposed algorithm to modify an inconsistent pair-wise weighting matrix in AHP	Ant Colony Optimization with AHP (ANTAHP)	1. Consistency ratio 2. Difference index	The proposed algorithm is hard to implement when the original opinion of decision-makers needs to be retained
Kou and Lin[103]	To evaluate the priority vector from PCM in AHP	Cosine maximization method (CM)	1. Euclidean distance 2. Minimum violation	The proposed approach is limited to derive priorities from complete and precise PCM only

based on gene expression[115] programming for the automatic clustering of data. GEP-cluster comprises clustering algebra as a new concept for identifying the best cluster with no prior knowledge and an automatic merging clustering algorithm for merging clusters automatically. The results indicated that the GEP-cluster algorithm is found noise sensitive as well as incompetent to high-dimensional datasets. Lee and Antonsson[116] introduced evolutionary strategy-based clustering, also called, ES-clustering for automatic clustering of data. In ES-clustering the initial population is encoded using genomes of variable-length and (10 + 60) ES selection strategy, and an improved mean square error is selected as a fitness function. Tseng and Yang[117] introduced an automatic clustering, called CLUSTERING, based on GA. In CLUSTERING, the data points are grouped into the small clusters using the nearest neighbor clustering method, then by GA, and using a difference between the WGS and BGS indices as the fitness function, all small clusters are combined to a larger cluster. Finally, a heuristic strategy is implemented to determine the cluster partition. The results revealed that the CLUSTERING algorithm outperformed when compared with K-means, complete-link, and single-link. Bandyopadhyay and Maulik[118] proposed a nonparametric VGA-clustering algorithm. In the VGA-clustering algorithm, a variable-length encoding is utilized for encoding chromosomes, composed of cluster coordinates, in the population, and the I-index is used to compute the fitness of each solution. The algorithm is evaluated on real-life datasets and produced an appropriate number of clusters. Bandyopadhyay and Saha[119] introduced a variable length and point-symmetry-based genetic clustering algorithm (VGAPS) for the automatic partitions of the dataset. The algorithm utilized point symmetry (PS)-based distance as a similarity measure and the Sym-index for fitness computation. The experimental analysis revealed that the VGAPSA produced better results. Liu et al.[35] presented an automatic clustering algorithm based on the GA, also called AGCUK. This algorithm aimed to determine cluster partition in the absence of the prior number of clusters. The algorithm adopted a real-code variable-length representation for encoding cluster centers in chromosomes and utilized DB index to compute the fitness of an individual solution. This algorithm produced better outcomes in terms of the misclassification rate. Ahmad and Dey[120] proposed K-means for the mixed (numeric and categorical) dataset and outperformed the K-means algorithm. K-harmonic mean with mixed data set was, also called KHMCMD, proposed by Ahmad and Hashmi[121] to cluster the mixed dataset. KHMCMD was found competent in producing satisfactory clustering outcomes in comparison to the K-means algorithm. Chang et al.[122] suggested an algorithm for automatic clustering. The algorithm adopted dynamic niching GA with niche migration to automatically determine the cluster centers and the number of clusters. The algorithm is compared with species-conserving GA (SCGA) and the dynamic fitness sharing (DFS)[122] algorithms using real-life datasets. It outperformed the compared algorithm in terms of determining a suitable number of clusters. A framework called ETSAs, based on tabu search, and genetic operators, is proposed by Pan and Cheng[123] for automatic clustering. The framework determines the optimal number of clusters utilizing an evolutionary approach to produce a competition population by utilizing two parallel reproduction processes. Based on the DE approach, an automatic clustering algorithm, and called ACDE, is proposed by Das et al. [43]. Each individual is encoded as a real value vector in the proposed algorithm, containing the activation threshold linked to the cluster centers. ACDE performed better in terms of classification error. Das et al.[44] proposed an automatic clustering algorithm based on a bacterial evolutionary algorithm called (ACBEA). The proposed algorithm consists of two operations: bacterial mutation and the gene transfer operation for population generation. These two operations are modified for managing the variable length in a chromosome that encodes cluster centers. Further, the fitness of each individual is computed using the CS index. Omran et al.[124] proposed a PSO-based dynamic clustering approach (DCPSO) for image[125, 126] segmentation. In the algorithm, the "best" number of clusters is selected using binary PSO. Then, the K-means is implemented to refine the chosen clusters. The results showed that DCPSO produced suitable cluster numbers. Cura[127] introduced a novel PSO clustering algorithm, called CPSO, that determines the number of clusters in both scenarios, whether the number of clusters is known, respectively. Further, two fitness functions are used in the algorithm: WGS index when the number of clusters is specified and the difference between the BGS and WGS indices when the number of clusters is unknown. The CPSO produced better results in terms of cluster quality. Chowdhury et al.[128] proposed an IWO-clustering for the automatic evolution of clusters. In this algorithm, the weed strings representing the population and encoding cluster centers are encoded using a variable-length encoding approach. Further, a modified Sym-Index is used as the fitness function in this algorithm. Kumar et al.[41] introduced a gravitational search-based automatic clustering algorithm for image segmentation. This algorithm adapted a variable-length approach for encoding the cluster centers to detect the number of clusters automatically. The algorithm produced optimally compacted and well-separated clusters. V. Kumar and D. Kumar[42] proposed an automatic clustering and feature selection algorithm based on the GSA. This algorithm aimed to produce an optimal number of clusters and to select relevant cluster features at run time. The algorithm utilized a modified I-index as fitness computation. The algorithm produced better results in terms of classification accuracy. A summary of the existing clustering methods is presented in Table 2.2.

Table 2.2: Summary of existing clustering methods

Author(s)	**Objective/Work Done**	**Technique proposed/Used**	**Performance Parameters**	**Research Gap(s) Identified**
Liu et al.[35]	Proposed an automatic clustering algorithm based on the genetic algorithm	Automatic genetic clustering for unknown K (AGCUK)	1. Degree of population diversity 2. Degree of selection pressure 3. Average and standard deviation of DB-index 4. Average and standard deviation of the number of clusters	The algorithm is limited to either numeric or categorical datasets only
Kumar et al.[41]	Proposed an automatic data clustering for image segmentation using gravitational search algorithm	Automatic clustering using gravitational search algorithm (ACGSA)	1. Within-cluster variation 2. Trace	The algorithm is limited to numeric attributes only
Kumar V. and Kumar D.[42]	Proposed an automatic data clustering and feature selection using the gravitational search algorithm (GSA)	Automatic clustering and feature selection using gravitational search algorithm (GSA_CFS)	1. Silhouette index 2. Classification error	The algorithm is limited to either numeric or categorical datasets only

Das et al.[43]	Proposed an automatic clustering algorithm using an improved differential evolution algorithm	Automatic clustering DE algorithm (ACDE)	1. Mean classification error 2. Standard deviation	The algorithm is limited to either numeric or categorical datasets only
Das et al.[44]	Proposed an automatic clustering algorithm based on the bacterial evolutionary approach	Automatic clustering using the bacterial evolutionary algorithm (ACBEA)	1. CS-measure 2. Inter-cluster distance 3. Intra-cluster distance	The algorithm is limited to either numeric or categorical datasets only
Sarkar et al.[113]	Proposed data clustering based on evolutionary programming	EP-clustering	DB-Index	This algorithm is sensitive to the initial cluster number, and limited to numeric attributes only
Chen et al.[114]	Proposed an automatic clustering algorithm based on gene expression programming	1. GEP-cluster algorithm 2. Automatic merging cluster algorithm	1. Mean square error 2. Success rate	This algorithm is sensitive to noise and found efficient for high-dimensional data
Lee and Antonsson[116]	Proposed a dynamic clustering algorithm based on evolutionary strategy approach	Evolutionary strategy-based clustering (ES-clustering)	Heuristic mean square error	The efficacy of the proposed algorithm is not tested on high-dimensional data
Tseng and Yang[117]	Proposed an automatic clustering algorithm using genetic algorithm and heuristic strategy	CLUSTERING	The average distance from the cluster center	The algorithm did not focus on mixed datasets

Bandyopadhyay and Saha[119]	Proposed a point symmetry-based genetic clustering algorithm for automatic partitioning of data	Variable-length point symmetry-based genetic clustering algorithm (VGAPS)	1. Minkowski score 2. Sym-index	This algorithm is focused on point-based symmetry only, and limited to either numeric or categorical data only
Ahmad and Dey[120]	Proposed K-mean clustering algorithm for the mixed data type	K-mean clustering for mixed dataset (KMCMD)	1. Micro-precision 2. Micro-recall	The cluster center initialization problem persists
Chang et al.[122]	Proposed an automatic clustering algorithm based on dynamic niching and niche migration	Dynamic niching and niche migration clustering DNNM clustering)	1. The average number of niches 2. Standard error	The algorithm is sensitive to the size of the initial population
Pan and Cheng[123]	Proposed a framework for automatic clustering algorithm based on tabu search and genetic operators	An evolution-based tabu search app-roach (ETSA)	1. Cluster validity index 2. (PBM-index) value	The algorithm is limited to either numeric or categorical datasets only
Cura[127]	Proposed a data clustering algorithm using enhanced particle swarm optimization (PSO)	Enhanced particle swarm optimization-based clustering (EPSO-clustering)	1. Error rate 2. Intra-cluster distance 3. Inter-cluster distance	Robustness decreases when dealing with problems where the numbers of clusters are unknown
Chowdhury et al.[128]	Proposed an automatic clustering based on invasive weed optimization (IWO) algorithm	IWO-clustering	Minkowski score	The algorithm is limited to either numeric or categorical datasets only

2.3 Optimal Surgical Team Selection

In a modern healthcare system, the provision of high-quality surgical services typically depends on symptoms of various patients. Classifying patients on the basis of their symptoms assist the decision-makers in identifying the target patient and making corresponding remedial decisions[129]. In recent years, several methods have been proposed by the authors for clustering of patients. One of the studies used k-means clustering algorithm to partition the patients based on their health status[129, 130]. Another study used a multilayer clustering approach for partitioning of Alzheimer disease patients into male and female groups[131]. In addition, few studies have used agglomerative hierarchical clustering for partitioning of patients[132] based on presence of comorbidities such as chronic pain & mental illness; obesity & mental illness; cancer; diabetes & renal disease[133]. A Bayesian nonparametric clustering approach was applied to divide patients having cancer into sub-groups to measure their anxiety and depression scores before psychotherapy[134]. Further, in a study authors proposed a hierarchical clustering algorithm incorporating genetic concept for partitioning the patients with or without depression[135]. Furthermore, many intelligent nature-inspired algorithms have also been proposed for data clustering. A study proposed a hybrid algorithm combining particle swarm optimization (PSO) and artificial bee colony (ABC) algorithms for data clustering[136]. Additionally, K-means (KMCMD)[120] and K-harmonic means (KHMCMD)[121] clustering algorithms have been used for performing clustering of mixed datasets. The functioning of most of these clustering algorithms are dependent on the predefined number of clusters. However, in real-life problems, for most of the datasets the number of clusters is not known beforehand.

Hence, the accurate estimation of an optimal number of clusters is a challenging task, and can affect the performance of a clustering algorithm also. Therefore, several algorithms, such as the gravitational search algorithm[41, 42], harmony search algorithm[137], and differential evolution algorithm[43], have been proposed for automatic clustering to address the aforementioned challenge. Automatic clustering algorithms require no prior information regarding the number of clusters. Instead, they evaluate the optimal number of clusters based on the dataset only. Furthermore, the quality of surgical care and the outcomes of patients depend on the performance of the surgical team. A recent study proposed a methodology to select a suitable surgical team based on patient characteristics[55]. The proposed method utilized k-prototypes algorithm for partitioning of patients and genetic algorithm (GA) for the selection of optimal surgical team. Although the study reported favorable outcomes, it focused only on the complication ratio for the surgical team selection. However, the success of a surgical procedure depends on various factors also such as a lower surgical readmission rate, lower mortality & complication rates, and higher patient satisfaction (surgical feedback) etc. Therefore, there is a need to attempt a modest contribution by focusing the feedback of patients who underwent surgery in the past along with complication rates to select an optimal surgical team. Further, the GA[138, 139] utilized by[55] is likely to experience premature convergence and diversity loss. In addition, Srinivas and Deb[140], Gu et al.[141], Hassanzadeh and Rouhani[142], Yuan et al.[143], and Nobahari et al.[144] have proposed several meta-heuristic approaches to prevent premature convergence. These algorithms have been found efficient in finding the optimal solution in a single computation. A summary of the existing work related to surgical decision-making is presented in Table 2.3.

Table 2.3: Summary of existing methods used in surgical decision-making

Author(s)	**Objective/Work Done**	**Technique proposed/used**	**Performance parameters**	**Research gap(s) Identified**
Gamberger et al.[131]	To partition patients with Alzheimer disease into homogeneous subgroups	A multilayer clustering approach	Clinical dementia rating score	The segmentation of patients with cognitive problems was not considered
Li et al.[134]	To partition patients with cancer into subgroups for measuring their anxiety and depression scores	Bayesian nonparametric clustering approach	Within cluster contrast	The proposed algorithm was sensitive to the initial sample size
Yu et al.[135]	To identify the subgroups of patients with and without depression	1. Hierarchical clustering based on the genetic concept 2. Distance matrix	1. Approximately unbiased probability value 2. Bootstrap probability value	This study considered the genetic data of a specific population only
Karthikeyan and Christopher[136]	Proposed hybrid PSO and ABC algorithm for data clustering	Hybrid PSO and ABC	1. Accuracy 2. Classification error percentage	The proposed algorithm focused on numeric attributes only
Ahmad and Hashmi[121]	Proposed k-harmonic mean for mixed data	K-Harmonic clustering algorithms	1. Intra-cluster distance 2. Inter-cluster distance	The proposed algorithm focused on numeric attributes only
Kumar et al.[137]	Proposed automatic data clustering using an adaptive harmony search algorithm	Adaptive harmony search algorithm	1. Inter-cluster distance 2. Intra-cluster distance 3. Trace	The proposed algorithm focused on numeric attributes only

Srinivas and Deb[140]	Proposed genetic algorithm based on non-dominant sorting for multi-objective optimization	Non-dominant sorting genetic algorithm	Chi-square test	The proposed algorithm exhibited a slow convergence rate
Gu et al.[141]	Proposed an evolutionary algorithm based on the projection of the current non-dominant solutions and equidistance interpolation for multi-objective optimization	Dynamic weight design method with MOEA/D	1. Benchmark functions 2. Mean 3. Standard deviation 4. Inverted generational distance	The algorithm lacked efficiency in solving higher-dimensional complex problems
Hassanzadeh and Rouhani[142]	Proposed a gravitational force-based algorithm for multi-objective optimization	Multi-objective gravitational search algorithm	1. Spacing metric 2. Generational distance metric	The algorithm experienced diversity loss in solving higher-dimensional complex problems
Yuan et al.[143]	Proposed a gravitational search algorithm based on strength Pareto for multi-objective optimization	Strength Pareto gravitational search	1. Convergence metric 2. Space metric 3. Generational distance metric 4. Diversity metric	Population diversity requires further improvement
Nobahari et.al.[144]	Proposed a multi-objective gravitational search algorithm based on non-dominant sorting for power transformer design	Non-dominant t sorting gravitational search algorithm	Normalized arithmetic mean	The algorithm experienced scalability loss while dealing with complex cases of power transformer design

Chapter 3

PROBLEM FORMULATION

In this chapter, research gaps, objectives, and methodology are discussed. The research gaps are identified from the existing literature related to patient prioritization, patients clustering, and surgical team selection. By considering these gaps, four research objectives were identified for healthcare resource optimization and scheduling. The main focus of the research is towards: patient prioritization from the surgical waiting list, postoperative surgical patients clustering, and optimal surgical team selection.

3.1 Research Gaps

While considerable interest has been found in healthcare resource management research over the past few decades, there are still numerous concerns that need to be addressed. Most of the techniques and algorithms proposed in the literature are based on mathematical and conventional decision-making methods in the literature. It is observed that soft computing capabilities are underexploited, and intelligence-based approaches are needed. There are the following research gaps which are identified:

1. Healthcare resources planning suffers from various issues like a poorly managed waiting list, disorganized health records, etc. Although different methods and techniques have been proposed in the literature, efficient resource scheduling and optimization still require more attention.

2. Most of the existing work has either been done to prioritize waiting lists or improve inconsistent matrices to derive consistent priorities. However, no work that focuses on improving inconsistent matrices and prioritizing waiting list together has been done so far. There is a need for a framework that can help achieve optimal and consistent priorities on the waiting list while improving inconsistent metrics.
3. Hospital management suffers from the issue of organizing an enormous number of patient medical records. Since patient medical records are unlabeled mixed data, there is a need to develop an efficient clustering approach to organize the patient medical records into optimal sub-groups.
4. Since surgical services play a crucial role in the healthcare system, the surgical team's performance significantly affects surgical outcomes. Patients always require quality surgical services, while hospitals are more concerned about selecting a suitable surgical team for patients. There is a need for a decision-making model to help the healthcare professional select a patient-centered optimal surgical team.
5. Meta-heuristic capabilities in healthcare need to be explored more. Hence, the relevant meta-heuristic-based algorithm and framework need to be developed to optimize and schedule healthcare resources.

3.2 Research Objectives

The objectives of the proposed work are given below:

(1) To study and explore various techniques and tools available for managing and analyzing healthcare data

(2) To propose a framework for scheduling and optimization of healthcare resources

(3) Design and Implementation of proposed framework using evolutionary algorithm(s)

(4) Testing and validation of the proposed work using real data.

3.3 Research Methodology

This research work aims at the development of a framework for scheduling and optimization of healthcare resources. Figure 3.1 demonstrates the flow of the adopted research methodology.

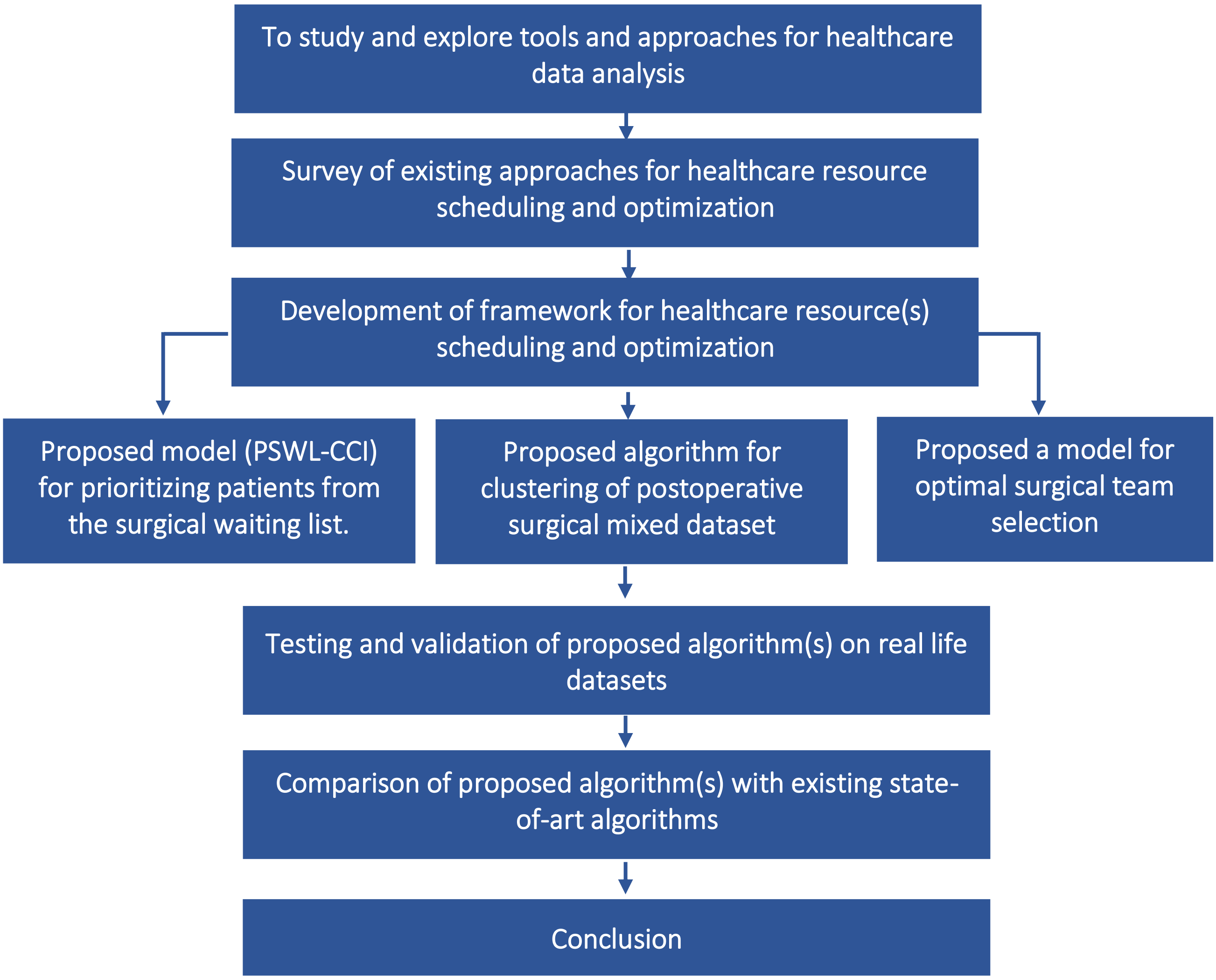


Figure 3.1: Flow of the Research Methodology

3.3.1 Research Methodology for Objective 1:

- Studied various approaches used for scheduling and optimizing healthcare resources, such as patient prioritization and selecting optimal surgical teams for a patient.
- Identified approaches available for finding patterns in healthcare resources, such as patient clustering using a mixed dataset.
- Identified the tools that can be used to analyze healthcare data.

3.3.2 Research Methodology for Objective 2:

- We have proposed a decision-making model PSWL-CCI for prioritizing patients from the surgical waiting list. Firstly, we have proposed a hybrid algorithm named HMWCA (Hybrid modified water cycle algorithm) for improving the cosine consistency index (CCI) of inconsistent pair-wise comparison matrices (PCMs), focusing on deriving consistent priorities of patients. The proposed algorithm utilizes the features of the evaporation-based water cycle algorithm (ER-WCA)[99], genetic algorithm, and 2-opt heuristic algorithm[145], where ER-WCA is modified with the concepts of salt concentration and infiltration. The proposed algorithm is tested on different inconsistent PCMs, and the proposed model is validated using a dataset obtained from a surgical department of a multispecialty hospital in India.
- Secondary, an efficient clustering algorithm for mixed datasets of postoperative surgical records is proposed. The proposed algorithm aims to group the mixed dataset into optimal partitions, and it focuses on a postoperative surgical dataset to partition the historical surgical patients into optimal subgroups. The algorithm

is tested on publicly available real-life datasets and compared with existing clustering algorithms. The algorithm is also validated using a surgical dataset obtained from a multispecialty hospital's surgical department.

- Thirdly, we proposed a model for surgical team selection. The proposed model involves two modules: surgical history management (SHM) module, i.e., arranging surgical patients into optimal surgical sub-groups, and surgical team selection (STS) module, i.e., selecting optimal surgical teams for a newly admit surgical patient based on the surgical sub-groups. In the SHM module, the earlier proposed clustering algorithm is modified and proposed to cluster surgical patients into optimal subgroups. In the STS module, a multi-objective optimization algorithm is proposed to select an optimal list of the surgical team for a given patient. The proposed algorithms are based on the artificial electric field algorithm[146]. The proposed algorithms are tested and validated on different publicly available real-life datasets and benchmark functions (SCH, FON, ZDT1)[147], and compared with existing algorithms. Further, the proposed algorithms are validated using the surgical dataset obtained from a multispecialty hospital's surgical department.

3.3.3 Research Methodology for Objective 3:

- The algorithms are implemented on Acer machines with 8GB of RAM, 1 TB hard disk, having Windows 10 OS installed. The experiments have been performed on MATLAB R2018a (9.4.0.813654) and R (3.5.1).
- The proposed PSWL-CCI model and its associated algorithms are implemented on MATLAB. The pair-wise comparison matrices (PCMs) are prepared based on

the surgical decision-makers ratings using questionnaires. The surgical priorities are computed using the analytic hierarchy process (AHP) and cosine maximization method (CM). The qualifying threshold for consistency of the cosine consistency index (CCI) is fixed to 0.90.

- For experimental analysis of the clustering algorithm, we have set the initial population size *(P)* to 20. The maximum numbers of clusters (K_{max}) is set to $\sqrt{size_{dataset}}$, and the Coulomb constant (K_0) is set to 500.
- In the surgical history management module (SHM), the initial population size (P_{SSIZE}) is fixed to 50. The maximum numbers of clusters (CLC_{max}) is set to $\sqrt{m}$ (Dataset size) and the Coulomb constant (K_0) is set to 500. For the surgical team selection (STS) module, patients' feedback rating is recorded using questionaries and telephonic conversation. In the proposed multi-objective optimization algorithm, the searching population size is fixed to 100. The external population size is fixed to 100 for SCH, FON, and ZDT1 functions. The initial value of the Coulomb constant is fixed to 500. The initial, final crossover probabilities are fixed to 1.0 and 0.0, respectively. The initial, final mutation probabilities are fixed to 0.01 and 0.001, respectively. The proposed algorithm is executed 100 times for SCH and FON functions and 250 times for ZDT1 function.

3.3.4 Research Methodology for Objective 4:

- The overall performance of the PSWL-CCI model is validated using a real dataset obtained from the surgical department of a multispecialty hospital in India. Subsequently, performance is compared with four existing prioritization methods in terms of Euclidean distance and minimum violation. Further, the proposed algorithm's performance is evaluated on six different inconsistent PCMs and compared with two existing approaches in terms of cosine consistency index (CCI), deviation score, and the number of iterations.

- The performance of the proposed clustering algorithm is evaluated on five real-life datasets and compared with six different existing clustering algorithms for mixed datasets in terms of average accuracy and standard deviation. Further, the proposed algorithm is also validated using a historical postoperative dataset obtained from a multispecialty hospital's surgical department.
- The performance of the surgical team selection model is evaluated on a real-life dataset obtained from a surgical department of a multispecialty hospital. Further, the proposed clustering is evaluated on nine real-life datasets and compared with existing non-mixed dataset clustering algorithm using three performance index[148], namely Silhouette Index (SI), Dunn Index (DI), and Davies-Bouldin index (DB). The proposed clustering algorithm is also compared with the mixed dataset clustering algorithm using two robustness measures: average accuracy and standard deviation. Subsequently, the proposed multi-objective optimization algorithm is evaluated on three benchmark functions, namely SCH, FON, and ZDT1 and compared with four existing multi-objective optimization algorithms in terms of mean of three performance metrics (CM metric[147], DM metric[147], and GD metric[149]).

Chapter 4

PRIORITIZING PATIENTS FROM THE SURGICAL WAITING LIST

The objective of this chapter is to propose a decision-making model to prioritize patients from the surgical waiting list. The model deals with two critical issues: First, to prioritize patients from the surgical waiting list. Second, to refine and optimize the cosine consistency index (CCI) of inconsistent pair-wise comparison matrix (PCM) to obtain consistent priorities. The cosine maximization method (CM), along with the analytic hierarchy process (AHP), is used to evaluate the resulting priority. To improve the inconsistent pair-wise comparison matrix (PCM), a novel hybrid algorithm HMWCA (Hybrid modified water cycle algorithm), is proposed and incorporated in the proposed model. The proposed algorithm is a hybrid algorithm that exploits the features of three traditional algorithms, namely the evaporation-based water cycle algorithm (ER-WCA), genetic algorithm, and 2-opt heuristic algorithm.

4.1 Decision-Making and Optimization Methods

This section briefly discusses a multi-criteria decision-making method, analytic hierarchy process (AHP), Cosine maximization method (CM), Cosine consistency index (CCI), Water cycle algorithm (WCA), Genetic algorithm, and 2-Opt heuristic algorithm.

4.1.1 Analytic Hierarchy Process (AHP)

AHP is a multi-criteria decision-making method, which helps multiple decision-makers in dealing with complex priority assessment problems. AHP decomposes a problem into a hierarchical structure of easily understandable sub-problems. These sub-problems can be studied individually. The hierarchical structure arranges the decision objective at the top, and each factor (criterion and corresponding sub-criterion) and the alternatives in descending order. The factors that lie at the same level are compared among themselves. The PCM is designed, and the weight (relative importance) of each factor is computed. The factor with the maximum weight is considered the best solution.

4.1.2 Cosine Maximization (CM) Method

The CM[103] method is introduced to derive the priority vector from the PCM in the AHP. CM method implemented an optimization model to derive reliable priorities. It maximizes the sum of the cosine angle between each value of a column in the PCM and the evaluated priority vector.

4.1.3 Cosine Consistency Index (CCI)

The consistency of PCM is a significant issue in AHP to derive a priority vector. The reliability of priority outcomes is highly dependent upon PCM's consistency. An inconsistent PCM leads to an unreliable priority. The cosine consistency index[103] (CCI) is introduced to measure the consistency of PCM. CCI is derived using an optimization model as follows:

$$CCI = \frac{C^*}{m} \tag{4.1}$$

Where, C^* and m are optimal objective function and matrix size, respectively, of the optimization model. Consistency of PCM is concluded from the value of CCI as follows:

$$CCI = 1, \text{ otherwise}, 0 < CCI < 1 \tag{4.2}$$

A PCM is entirely consistent if the value $C^* = m$, resulting in CCI = 1. Otherwise, it is considered as near consistent if 0 < CCI < 1. Here, 0.90[103] is considered as the minimum value to satisfy consistency (CCI).

4.1.4 Water Cycle Algorithm (WCA)

The WCA[99] mimics the water cycle process in nature, where rivers and streams flow towards the sea. WCA begins with the process of rainfall, where individual raindrop generates streams. Then, it selects some of the good streams as rivers and the rest considered as tributary streams from the generated streams. Finally, it chooses the best individual stream as the sea. The steps of WCA are as follows:

A. Initialization

At first, it randomly generates an initial population of streams N_{pop}. Then, it selects some of the best individual streams N_{sr} as sea and rivers. Finally, it computes the rest of the streams N_{stream} as follows:

$$N_{sr} = \text{River} + \underbrace{1}_{\text{sea}}, N_{stream} = N_{pop} - N_{sr} \tag{4.3}$$

Depending on the rate of water flow, there is a significant variation in the volume of water entering a river or the sea. WCA computes the tributary streams for each river or the sea. WCA computes the tributary streams for each river and sea in the following way:

$$NS_n = round\left\{\left|\frac{Cost_n - Cost_{N_{sr}+1}}{\sum_{n=1}^{N_{Sr}} C_n}\right| * N_{stream}\right\} \tag{4.4}$$

Where, $n = 1, 2, 3 \dots . N_{Sr}$, and $Cost_n$ represents the objective function.

B. Flow and Exchange

A stream flows to a river or sea using a randomly chosen distance and updates new position as follows:

$$\mathrm{x} \in (0, c * d), \text{ where c} = 2 \tag{4.5}$$

Where, d is the distance between stream and river/sea. The value of c, greater than one, enables streams to flow in different directions toward the rivers and sea.

$$X_{stream}^{i+1} = X_{stream}^{i} + rand * C * \left(\frac{X_{river}}{sea^{i}} - X_{stream}^{i}\right) \tag{4.6}$$

$$X_{river}^{i+1} = X_{river}^{i} + rand * C * \left(X_{sea}^{i} - X_{river}^{i}\right) \tag{4.7}$$

The stream exchanges its position with the river when the solution representing a stream is better than its connecting river. Similarly, the river exchanges its position with the sea when the solution representing a river is better than the sea.

C. Evaporation and Rainfall Process

Evaporation causes seawater to vaporize as rivers and streams flow into the sea, and initiate a new rainfall. The evaporation operator is applied on both rivers and streams that flow into the sea and streams flowing towards rivers. The distance (closeness) from the sea computes evaporation between the sea and river and between sea and stream. The number of tributary streams NS_n flowing towards a river, called the river's quality, computes the evaporation between rivers and streams. More tributary streams flowing towards a river increase the quality of the river and result in a slower evaporation rate. Rivers with poor quality comprise of lesser number of tributary streams. All such types of rivers cannot decrease their distance d_{max} to sea, and they tend to evaporate at a fast rate following some movements. d_{max} determines the search intensity near the sea, which is the best-obtained solution. A small value of d_{max} stimulate the search intensity near the sea while a higher value for d_{max} inhibits extra searches. Hence, for search exploration, the d_{max} is decremented as follows:

$$d_{max_{i+1}} = d_{\mathrm{max}_i} - \frac{d_{\mathrm{max}_i}}{Maximum\ Iteration} \tag{4.8}$$

4.1.5 Genetic Algorithm

Genetic algorithm (GA), a population-based meta-heuristic, mimics the process of biological evolution known as natural selection. It is a process that selects the best fit individuals to give birth to the offspring of the next generation. It ensures to derive the optimal solutions in time. Algorithm 1 describes the process of GA.

4.1.6 2-Opt Heuristic Algorithm

2-Opt[145], a local search algorithm, proposed to solve the traveling salesman problem. 2-Opt algorithm increases the convergence speed and enhances neighborhood search, which helps in finding the optimal solution. It begins with an initial network (tour).

Then it separates two edges (Figure 4.1) from the network and reconnects the two edges in several different ways. This entire process occurs several times for a different set of two connected edges. Finally, it computes the length of each generated tour and chooses the optimized tour.

Algorithm 1: Genetic Algorithm

```
Input: Initialize Population Size, Crossover Rate, Mutation Rate, Maximum Iteration
Output: Optimal solution
1. Randomly create an initial solution of population size
2. Calculate the fitness of each solution
3. while $M_I$ <= Maximum Iteration do
   3.1 for $k$ =1: Population Size do
       (a) Randomly selection two best-fit solutions from the population
       (b) Apply cross over on selected solutions and create new solutions
       (c) Apply mutation to new solutions
       (d) Evaluate the individual fitness of mutated solutions
       (e) Replace the least-fit population with new solutions and update the
           population
       (f) $k = k + 1$
       end for
   3.2 $M_I = M_I + 1$
   end while
```

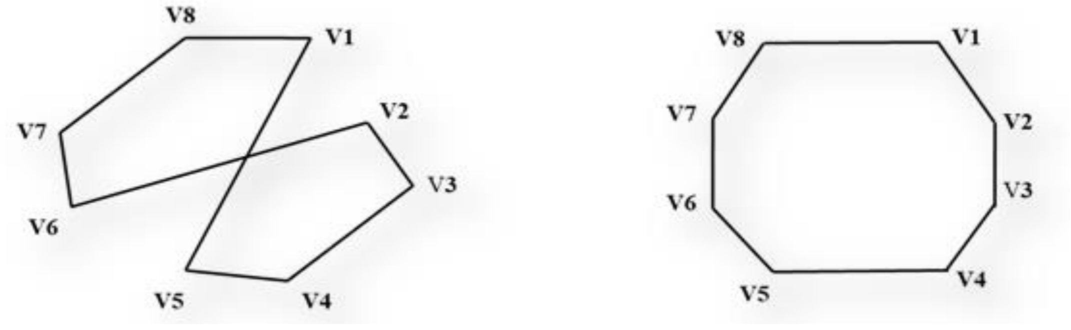


Figure 4.1: A 2-Opt move: original tour on the left and the resulting tour on the right

4.2 Proposed Prioritization Model

This section describes the proposed model, various phases, and methods used in detail. The first phase selects the factors related to the patient and generates the hierarchical structure containing these factors. The second phase compares all factors and generates PCM by assigning weights to factors based on the unanimous decision of all experts generated. The third phase evaluates the consistency level of each PCM and validates it against the threshold value. The fourth phase performs two steps; modification of inconsistence PCM using the proposed hybrid algorithm (HMWCA) and derivation of patients' priorities. Figure 4.2 represents the detailed workflow of the proposed model.

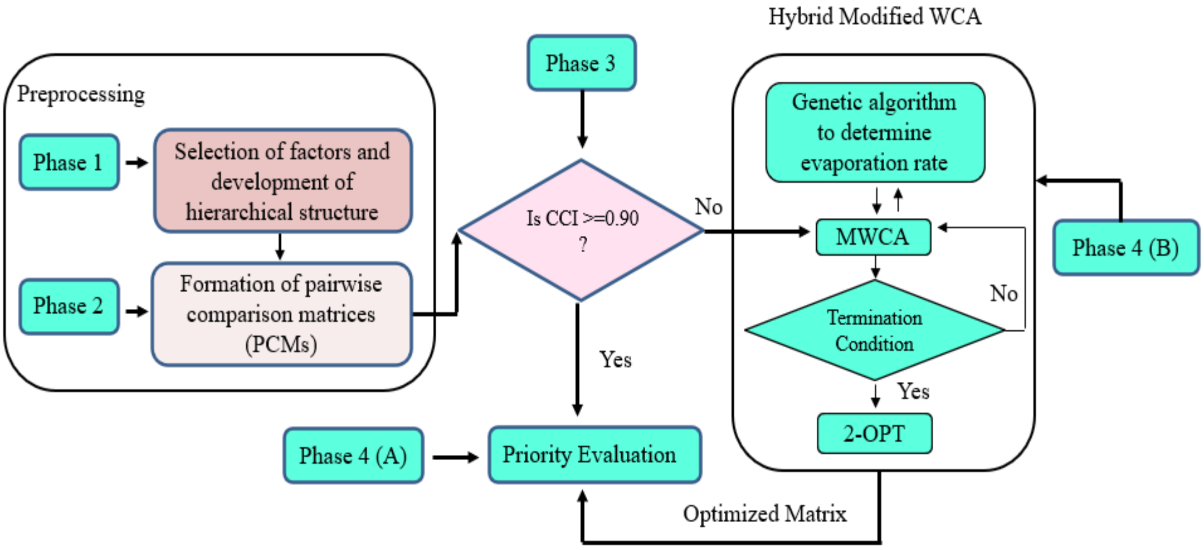


Figure 4.2: Model for prioritizing surgical patients

4.2.1 Phase 1 - Selection of Factors and Development of Hierarchical Structure

To assess the priority of patients, the literature[72] proposed eight significant factors (criteria). Following the selection of criteria, another researcher[24] proposed six additional risk-based sub-criteria. Therefore, in this phase, an overall eight criteria and six sub-criteria are considered necessary. Six patients are considered as alternatives to evaluate the priority of patients for surgery. All of these factors guide experts in the collective decision-making process. An explanation of all the criteria (4.1.1.– 4.1.8.) and sub-criteria (a–f) is as follows:

A. Disease Severity (DS)

It represents the physiological decompensation or deterioration of the organs of a patient. It is related to the health of a patient after he/she has contracted a disease. Severity can be minor, moderate, major, and extreme.

B. Pain

Pain is an unpleasant sensory and emotional experience associated with actual or potential tissue damage or described in the context of such harm. Pain may be mild, steady, or severe.

C. Rate of Disease Progression (RDP)

It represents the change in how the disease affects a patient as it progresses from its initial stages to its peak and then to its resolution. It relies on age, health state, and type of disease and thus varies from patient to patient. This criterion further has six sub-criteria:

(i) **Risk of Death (RD):** This represents the risk of mortality while waiting for surgery. Different risk levels exist for different patients waiting for the same surgery. This sub-criterion ensures the higher priority of medical attention to the patient at a higher risk of death.

(ii) **Risk of Deterioration of Disease Intensity (RDI):** It represents the possibility of patients developing additional health issues or severe complications while waiting for surgery.

(iii) **Risk of Deficiency in Surgery Effectiveness (DES):** It represents the possibility that a delay in surgery can reduce its effectiveness.

(iv) **Risk of Past Complications (RPCs):** It represents the risk associated with the outcome of surgery due to a patient's medical history.

(v) **Risk of Affecting Adjacent Organs (RAO):** This represents the threat caused by the spreading of the disease from the organ in concern to an adjacent organ.

(vi) **Progression that Might Affect Survival (PAS):** It represents the possibility that a delay in surgery may affect a patient's chances of survival.

D. Difficult in Doing Activities (DDA)

It represents the difficulty faced by a patient in performing daily activities that they could once perform smoothly before the onset of disease.

E. Probability and Degree of Improvement (PDI)

This represents the expected benefit of surgery in improving a patient's health and quality of life.

F. Time on Waiting List (TWL)

It represents the time between the registration of the patient for surgery and the actual surgery is performed. This criterion has importance from a patient's perspective and helps patients form a perception of the quality of service they receive.

G. Limitation in Caring Dependents of Someone (LCDS)

It is defined as the limitation in a patient's ability to look after their dependents.

H. Limitation to Work/Study/Seek Employment (LSE)

This represents a patient's hindrance to work, study, or seek a job due to their health.

4.2.2 Phase 2 - Formation of Pair-wise Comparison Matrices (PCMs)

This phase compares different factors (criteria and sub-criteria) among themselves. A questionnaire (formulated using the AHP method) is sent to experts. The experts are then asked to assign a weight for every comparison using linguistic variables. Subsequently, the linguistic variables are translated into numbers using Table 4.1. Finally, PCMs are designed by integrating the weights provided by all the experts for each comparison. The steps followed in this phase are as follows:

(a) Each criterion is compared with another criterion, and then PCM is formed by determining an appropriate weight for each comparison.

(b) Each sub-criterion is compared to another sub-criterion, and then PCM is formed by determining an appropriate weight for each comparison.

A day before surgery, a list of the first six patients from the waiting list is provided to the experts. Then, the patients are compared with one another in terms of each criterion and sub-criterion. As a result, three PCMs: (a) Criteria PCM, (b) Sub criteria PCM, and (c) Patients PCM are formed by determining an appropriate weight for each comparison. To assign weights to the criteria, sub-criteria, and patients, experts are provided with a linguistic scale of numbers, which indicates how significant or assertive one factor is over another.

Table 4.1: Linguistic scale and its description

Linguistic expression	**Explanation**	**Importance level**
Uniformly important	The contribution of both tasks is the same	1
Uniformly to averagely important		2
Averagely important	One task is preferred slightly over another	3
Averagely to emphatically important		4
Emphatically important	One task is preferred emphatically over another	5
Emphatically to very emphatically important		6
Very emphatically important	One task is preferred very emphatically over another	7
Very emphatically to most important,		8
Most important	One task is considered to be exceedingly important in comparison to another	9
If a non-zero number from importance level is assigned to the comparison between task p and task q, a reciprocal value is assigned to the comparison between task q and task p		The reciprocal value of above

4.2.3 Phase 3 - Consistency Measurement of Pair-wise Comparison Matrices (PCMs)

As mentioned above, consistency measurement is an essential process in the AHP, and a derived priority vector fails to satisfy reliability until its PCM shows consistency.

Therefore, to measure the consistency of each PCM, the following steps are used:

(a) For each PCM, Eq. (1) computes the value of CCI.

(b) Based on the value of CCI, Eq. (2) determines the consistency level of each PCM.

4.2.4 Phase 4 - Priority Evaluation of Surgical Patients

Once the consistency of PCM's is measured, this phase is implemented. This phase performs two activities.

A. Evaluation of Patients' Priorities

This section computes the priority of patients if PCMs meet the defined threshold. A rank is assigned to each patient representing their position in the final waiting list. The steps to compute priority are as follows:

(a) Priority weights of all factors and patients concerning each factor are computed.

(b) Each sub-criterion weight (SC_{w_i}) is multiplied by its parent criterion weight (P_{cw}) to give corresponding resulting global priority weight of factors.

$$SC_{w_i} = SC_{w_i} * P_{cw} \tag{4.9}$$

All factors with priority weight are arranged in columns. Priority weights of each patient corresponding to their factor are arranged in rows.

(c) Each criterion/sub-criterion weight is multiplied by the patient's weight under it. After the integration of all weights, the final priority vector of patients is computed. For this, we have:

$$Patient_i = \sum_{j=1}^{m} Patient_i * Criteria_j \tag{4.10}$$

Where, i = 1, 2, 3…number of patients, and j = 1,2...number of criteria.

Finally, according to the computed priority vector values, patients are assigned a rank representing the order of the treatment.

B. Modification in Inconsistent PCM using Proposed Hybrid Algorithm

This section is invoked when PCM fails to satisfy the threshold of consistency. Unlike AHP, in which the unavailability of decision-makers to modify PCM results in unreliable priorities, the proposed algorithm automatically modifies the inconsistent PCM on an iterative basis. This phase implements our proposed algorithm, HMWCA. The proposed algorithm integrates GA, WCA, and 2-OPT heuristic algorithms to modify inconsistent PCM and achieves the optimally consistent matrix. An optimally consistent matrix results in CCI with a satisfying threshold and the least deviation between modified PCM and original PCM. Algorithm 5 shows the steps followed in the proposed algorithm.

(i) Encoding PCM Elements for WCA algorithm

WCA solving TSP[100] generated the population by constructing the tours (positions) as the permutation of numbers representing the nodes, e.g., V = V_{Source}, V_2,…$V_{Destination}$, V_{Source} visited. In our proposed algorithm, nodes are set based on the values of elements of the PCM. As the elements of PCM are reciprocal a_{ij}= 1/a_{ji} , it can encode the lower triangular elements of the matrix as nodes[91]. Matrix A, with order $m = 4$, is given in Figure 4.3. The number of elements of $Geno_A$ can then be determined by $(m^2 - m)/$ 2. If matrix A fails to satisfy consistency, the corresponding sequence nodes are considered as “inconsistent-tour.” To obtain a “consistent-tour,” the new scale value of each element replaces the original matrix's old values. The new values are generated by fractionalizing from the original values into several candidate values. If the original element has a scale value between 2 and 9, it gets fractionated into candidate values between 1 and 9. Otherwise, if the element has a scale value between 1/9 and 1/2, it gets fractionated into candidate values between 1/2 and 1/9. In case the original element is 1, the value remains unchanged. Each candidate value is built from the specified scale value using the fraction factor. Suppose $Geno_A$ is the original value on node A and m is the matrix size; then, $Geno_A$ is described as:

$$Geno_A = ge_1, ge_2, \ldots \ldots ge_{\frac{m^2-m}{2}} \tag{4.11}$$

Each element of $Geno_A$ is fractionated into several candidate elements ge_{ri}. Value i is the index of the candidate element.

$$\begin{cases} ge_{ri-1} + \varphi, & \text{if } 1 < ge_r \leq 9 \\ \frac{1}{\frac{1}{ge_{ri}} - 1} + \varphi, & \text{if } \frac{1}{9} < ge_r \leq 1 \end{cases} \tag{4.12}$$

Where, ge_{r0} = 1, r = 1, 2, …$\frac{m^2-m}{2}$ and i= 1, 2, …$n - \frac{1}{\varphi}$. The value of φ is set to 1.

$$A = \begin{bmatrix} 1 & a_{12} & a_{13} & a_{14} \\ a_{21} & 1 & a_{23} & a_{24} \\ a_{31} & a_{32} & 1 & a_{34} \\ a_{41} & a_{42} & a_{43} & 1 \end{bmatrix} \qquad Geno_A = a_{21} - a_{31} - a_{32} - a_{41} - a_{42} - a_{43}$$

(a) Pairwise comparison matrix (PCM) $'A'$

(b) Encoding of PCM $'A'$ into genotype

Figure 4.3: PCM ‘A’ (with size $m = 4$) and its encoding in genotype

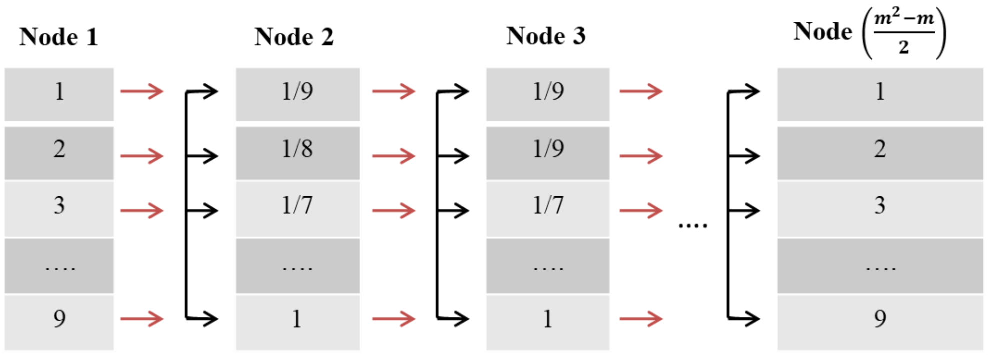


Figure 4.4: Candidate in each node from the PCM $'A'$

Figure 4.4 shows the nodes and their elements. Each element of each node has a combination with the element of other nodes.

(ii) Objective Function

This section focuses on improvement of inconsistent PCM. The first component considered is the cosine consistency index (CCI). As a consistent PCM satisfies $\text{CCI} \geq 0.90$ and for a higher value of CCI, PCM becomes more consistent, the first objective is to maximize CCI than the threshold 0.90. Further, the second objective here is to minimize the distance (deviation) between the original inconsistent and improved consistent PCM to maintain experts' original judgment. Euclidean distance is used to measure the deviation (DI) between two matrices. A smaller DI indicates higher similarity between the two PCM’s. Finally, PCM is improved and optimized in such a manner that the improved PCM satisfies the CCI threshold (0.90), and its deviation is minimum from the original PCM. So, the objective function is computed as:

$$F(x) = \frac{1}{1 + G_1(x)} + G_2(x) \tag{4.13}$$

Where, $F(x)$ represents the deviation score. $G_1(x)$ and $G_2(x)$ are CCI of improved

PCM 'x' and deviation of PCM x and original PCM. Algorithm 2 presents the sequence of steps followed to obtain an optimal solution.

Algorithm 2: Selection of Optimal Solution

```
Input: Population Size, Consistency Index (CCI), Minimum deviation score
Output: Optimal solution
1. for $i$=1: Population Size do
   1.1 Compute the following:
       (a) Number of genotypes satisfying CCI threshold (GSC): Genotypes with CCI >= .90
       (b) Number of genotypes dissatisfies CCI threshold (GDC): Genotypes with CCI < .90
   1.2 if GSC == 0 or GSC == 1 then
          Sort the genotypes in descending order of CCI
       else
          Sort the genotypes in ascending order of minimum deviation score
       end if
   1.3 $i = i + 1$
   end for
```

Here, $GSC = 0$ represents that no solution meets the CCI threshold, and $GSC = 1$ means only one solution meets the CCI threshold.

(iii) Modified Evaporation Rate of WCA (MWCA)

The proposed HMWC algorithm (Figure 4.5) considers two significant aspects of moving water bodies to modify the evaporation process. The first aspect is concerned with salt content in moving water bodies. Our proposed algorithm assumes that the rate of evaporation of the river depends on the amount of salt that the river carries, which in turn depends on the volume of streams. The second aspect is concerned with infiltration.

It affects the volume of water (called surface run-off) flowing over the land surface. An increase in infiltration decreases surface run-off. The greater the salt content, the higher the solution's mass and density, and the lesser the evaporation rate. The process of modified evaporation rate is as follows:

$$S_{alt} = \mathrm{LB} + (\mathrm{UB} - \mathrm{LB}) * \mathrm{rand}() \tag{4.14}$$

Here, S_{alt} represents the salt percentage. The upper bound (UB) for the sea is set to 3.5, whereas the lower bound (LB) is fixed to 0.05. For river salt value UB and LB are set to the salt value of sea and 0.05. For stream salt value, UB and LB are set to the river's salt value and 0.01, respectively.

$$\begin{cases} M_i = S_{alt_i} * V_i, \quad \text{Where,} \quad V_i = \frac{CCI_i}{DI_i} \\ DI_i \propto M_i, \quad i = \{Sea, River_{1,2,.....Nsr-1}, Stream_{1,2,3...N_{stream}}\} \end{cases} \tag{4.15}$$

In this equation, M, S_{alt}, V, DI represent mass, volume, salt value, and density of rivers, streams & sea. $V = \mathrm{CCI/DI}$ represent infiltration (volume of water). The value of $\mathrm{CCI/DI}$ varies with the change in position of the updated solution.

(i) The evaporation for each river is computed in the following way:

$$TD_{river_i} = DI_{river_i} + \sum_{s=1}^{N_{Stream}} DI_{Stream_i}, ER_{river_i} = \frac{1}{TD_{river_i}} \tag{4.16}$$

(ii) The evaporation rate, modified after the movement of streams into the river, is as follows:

$$ER_{river_i} = \frac{1}{TD_{river_i}} \tag{4.17}$$

Here, TD represents the total density of the river. ER represents the initial evaporation rate of the river whereas ER_m represents the modified evaporation rate of the river. Each river's evaporation rate is updated to a new position as all streams update their positions towards their rivers, and rivers update their position towards the sea. The new evaporation rate corresponding to the updated position is compared with the previous position, and rivers with a higher evaporation rate evaporate. The steps followed to compute the modified evaporation rate are as follows:

(i) When evaporation between river and streams occurs:

$$TD_{river_i} = DI_{river_i} + \frac{1}{ER_{river_i}}, ER_{river_i} = \frac{1}{TD_{river_i}} \tag{4.18}$$

(ii) When evaporation does not occur:

$$ER_{river_i} = ER_{m_{river_i}} \text{ where, } i = 1,2,3 \ldots . N_{sr} - 1 \tag{4.19}$$

(iv) Genetic Algorithm with MWCA (GMWCA)

To improve the solution's quality, we implemented the genetic algorithm with the mutation operator and single-point crossover on the streams flowing towards their river. As a result, an optimal solution is obtained, which helps MWCA to compute the evaporation rate of the river. Algorithm 3 shows the steps followed in the genetic algorithm.

(v) 2-OPT Algorithm with GMWCA

The 2-OPT algorithm improves the genotype (PCM element) obtained using Algorithm 3. It presumes that the genotype obtained using Algorithm 3 lacks optimality due to falling in local minima. Algorithm 4 shows the steps to be followed in the 2-OPT algorithm.

Algorithm 3: Genetic Algorithm with MWCA (GMWCA)

Input: Genotype representing the river ($Geno_{river}$) and tributary stream ($Geno_{stream}$), Density and CCI of $Geno_{river}$, Deviation score ($F(x)$) of $Geno_{river}$

Output: Optimal solution with density

```
1. Generate new $Geno_{stream(s)}$ by applying cross over and mutation between $Geno_{river}$ and $Geno_{stream(s)}$ solutions
2. Compute CCI and $F(x)$ of new $Geno_{stream(s)}$using Eq. (4.13)
3. Select the best $Geno_{stream}$ having minimum $F(x)$ among $Geno_{stream(s)}$
4. if CCI $_{Geno_{stream}} \geq CCI_{Geno_{river}}$ && $F(x)_{Geno_{stream}} < F(x)_{Geno_{river}}$ then
    4.1 Calculate the density of $Geno_{stream}$
    4.2 Return $Geno_{stream}$ with its density.
   else
       Return $Geno_{river}$ as optimal solution
   end if
```

Algorithm 4: 2-Opt Algorithm with GMWCA

```
Input: Genotype representing sea ($Geno_{sea}$)
Output: Optimal genotype (PCM element)
1. Compute CCI, DI and objective function $F(x)$ for $Geno_{sea}$
2. Initialize a TSP tour ($Geno_{tsp}$) for $Geno_{sea}$ and calculate the length of the
   tour ($Geno_{TL}$)
3. for $m$ =1: $Max\_Iteration$ do
   3.1 for each $i$=1 in $Geno_{TL}$-2 do
           for j= $i + 2$ in $Geno_{TL}$ do
           (a) Construct the population ($Pop_{geno}$) of $Geno_{tsp}$ using the
               following steps:
               (i) Swapping the position of $i^{th}$ element with $j^{th}$ element,
                   and position of $(i + 1)^{th}$element with $(j + 1)^{th}$ element
              (ii) Compute CCI, DI, and $F(x)$ for $Geno_{tsp}$ of using Eq.
                   (4.13)
           (b) j = j + 1
           end for
       end for
   3.2 for each ($Geno_{tsp}$) in $Pop_{geno}$ do
           if $CCI_{Geno_{tsp_i}}$ >= 0.90 && $F(x)_{Geno_{tsp_i}} < F(x)_{Geno_{sea}}$ then
               Replace the $Geno_{sea}$ with $Geno_{tsp_i}$
           end if
       end for
   3.3 Replace the $Geno_{tsp_i}$ with $Geno_{sea}$
   3.4 $m = m + 1$
   end for
```

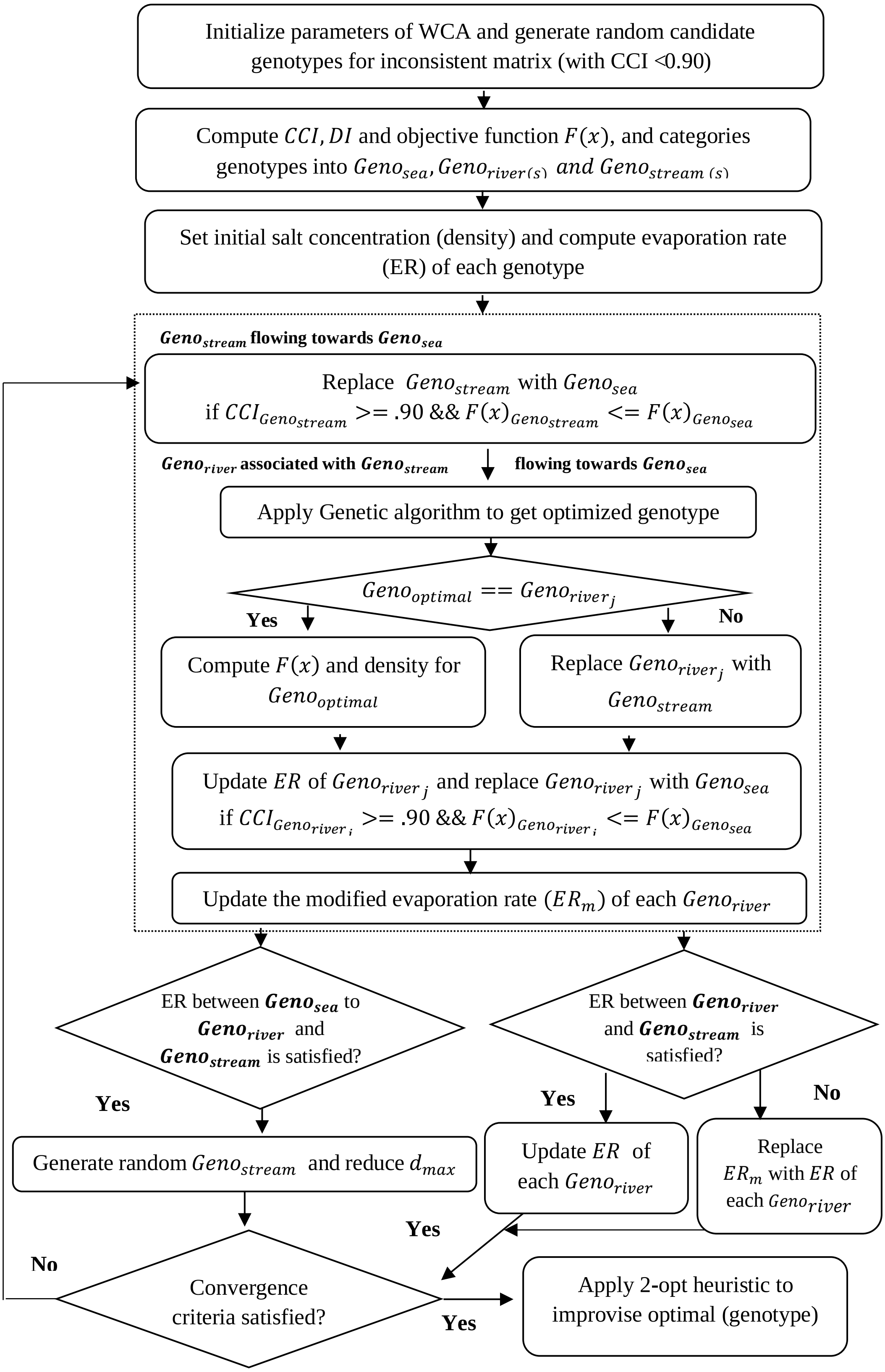


Figure 4.5: Flowchart of the proposed hybrid algorithm (HMWCA)

Algorithm 5: Proposed Hybrid Modified Water Cycle Algorithm (HMWCA)

```
Input: Inconsistent pair-wise comparison matrix (PCM), Population size ($N_{pop}$), No. of rivers and sea ($N_{sr}$), No. of streams ($N_{sn}$), Maximum Iterations ($MaxIt$), and $d_{max}$
Output: Optimal matrix elements (genotype)
1. Randomly generate the initial population of streams (genotypes), and compute $CCI$, $DI$ & objective function $F(x)$ using Eq. (4.11)
2. Sort the genotypes using Algorithm 2 and categories into $Geno_{sea}$, $Geno_{river(s)}$, and $Geno_{stream(s)}$
3. Set initial salt density ($TD$) of $Geno_{sea}$, each $Geno_{river}$ and each $Geno_{stream}$ using Eqs. (4.14) - (4.16)
4. Compute the initial evaporation rate ($ER$) for each $Geno_{river}$ using Eq. (4.16)

5. while $M_i <= MaxIt$ do
   5.1 for each $Geno_{stream}$ moving to $Geno_{sea}$ do
         Update the position of $Geno_{stream}$ using Eq. (4.6), and then compute
         $F(x)$ using Eq. (4.13)
             if $CCI_{Geno_{stream}} >= 0.90$ && $F(x)_{Geno_{stream}} <= F(x)_{Geno_{sea}}$ then
                 Replace the $Geno_{stream}$ with $Geno_{sea}$
             end if
         end for

   5.2 for each $Geno_{river}$ in $N_{pop}$ do
         (a) for each $Geno_{stream}$ moving to $Geno_{river\ j}$ do
               Update the position of $Geno_{stream}$ using Eq. (4.6) and compute optimal genotype and its $TD$ using Algorithm 3
                 if $Geno_{optimal}$==$Geno_{river\ j}$ then
                 (i) Compute $F(x)$ for $Geno_{optimal}$ using Eq. (4.13)
                 (ii) Update $TD$ of new $Geno_{optmial}$ and add it to the total $TD$ of $Geno_{river\ j}$
                 else
```

```
            Replace the $Geno_{river\ j}$ with new $Geno_{stream}$ and update the total
            density of $Geno_{river\ j}$
              if $CCI_{Geno_{river_j}} >= .90\ \&\& F(x)_{Geno_{river_j}} <= F(x)_{Geno_{sea}}$ then
              Replace $Geno_{river\ j}$ with $Geno_{sea}$ and update the total $TD$ of
              $Geno_{river\ j}$
              end if
            end if
          end for
      (b) Update the modified evaporation rate ($ER_m$) of each $Geno_{river}$ using Eq.
          (4.17)
      end for

5.3   for each $Geno_{river}$ moving to $Geno_{sea}$ do
        Update the position of $Geno_{river}$ using Eq. (4.7) and compute $F(x)$ for
        $Geno_{river}$ using Eq. (4.13)
          if $CCI_{Geno_{river_j}} >= 0.90\ \&\&\ F(x)_{Geno_{river_j}} <= F(x)_{Geno_{sea}}$ then
            Replace $Geno_{river_j}$ with $Geno_{sea}$ and update the $ER_m$ of $Geno_{river_j}$
            using Eq. (4.17)
          end if
        end for

5.4   for each $Geno_{stream}$ flowing to $Geno_{sea}$ do                  //evaporation
        if $norm\left(F(x)_{Geno_{stream}} - F(x)_{Geno_{sea}}\right) < d_{max}\ or\ (rand < 0.1)$ then
          Generate new $Geno_{stream(s)}$
        end if
      end for

5.5   for each $Geno_{river}$ flowing to $Geno_{sea}$ do                   //evaporation
        if $norm(F(x)_{Geno_{river}} - F(x)_{Geno_{sea}}) < d_{max}\ or\ (rand < 0.1)$ then
          Generate new $Geno_{stream(s)}$
        end if
      end for
```

```
    5.6 for each $Geno_{stream}$ flowing to $Geno_{river}$ do                //evaporation
            if $e^{\frac{M_i}{MaxIt}}$ $and\ ER_{mi} > ER_i$ then
            Generate new $Geno_{stream(s)}$ and update the $ER$ of $Geno_{river}$ using
            Eq. (4.18)
            else
            Replace $ER_m$ with $ER$ using Eq. (4.19)
        end for
    5.7 Reduce the value of $d_{max}$ using Eq. (4.8)
    5.8 Sort the genotypes using Algorithm 2
    5.9 $M_i = M_i + 1$
    end while
6. Compute optimal genotype using Algorithm 4.
```

4.3 Experimental Evaluation of Proposed Prioritization Model

This section discusses a case study of a multi-specialty hospital to prioritize surgical patients using the proposed prioritization model, PSWL-CCI. Subsequently, the performance of the proposed hybrid algorithm is evaluated by solving some inconsistent comparison matrices and comparing it with existing consistency improvement methods. Finally, the performance of PSWL-CCI is compared with existing prioritization methods. The multi-specialty hospital, which is the focus of our case study, is Shri Mahant Indiresh Hospital, Dehradun, Uttarakhand, India. This hospital is associated with Shri Guru Ram Rai (PG) Institute of Medical and Health Sciences in Dehradun, India. PSWL-CCI is implemented using datasets from the hospital’s orthopaedic surgery department. The hospital has 11 wards and 1000 beds, out of which approximately 90 beds are reserved for the orthopaedic department. In contrast, the remaining are reserved for other departments such as medicine, surgery, and gynaecology. There are 7 ICUs, each having approximately 10 beds and 17 air-conditioned fully modular operation theatres in the main operation section and 1 operation theatre in the emergency section. Every day, approximately 2000 patients are treated here, and every month approximately 2000 major surgeries are carried out at the main section of the operation theatre. The orthopaedic department comprises 3 neurosurgeons, 20 orthopaedic surgeons, and 10 anaesthetics. Approximately 1500 other employees in the hospital provide technical and support services. The symbols used in the proposed model are presented in Table 4.2.

Table 4.2: Symbols used in proposed PSWL-CCI model

Symbol	**Definition**
A	A Pair-wise comparison matrix (PCM)
A'	Repaired PCM
$F(x)$	Objective fitness function
ge	Candidate elements of genotype
S_{alt}	Salt percentage of rivers, streams, and sea
m	Matrix Size ($m * m$)
$Geno_A$	Genotype (row vectors comprising the lower triangular element of PCM A)
DI_i	Deviation (distance) between two genotypes (PCM elements)
M	Mass of rivers, streams, and sea
V	The volume of rivers, streams, and sea
C^*	The optimal objective function value
CCI	Cosine consistency index
TD	The total density of the river
ER	Initial evaporation rate of the rivers
ER_m	Modified evaporation rate of the rivers

4.3.1 Phase 1 - Selection of Factors and Development of Hierarchical Structure

The factors (Criteria and associated sub-criteria) are selected by unanimous decision of decision-makers (experts). The selected factors are converged into a hierarchical structure (Figure 4.6) using a multi-criteria decision-making method, AHP, to assess the patients' priority.

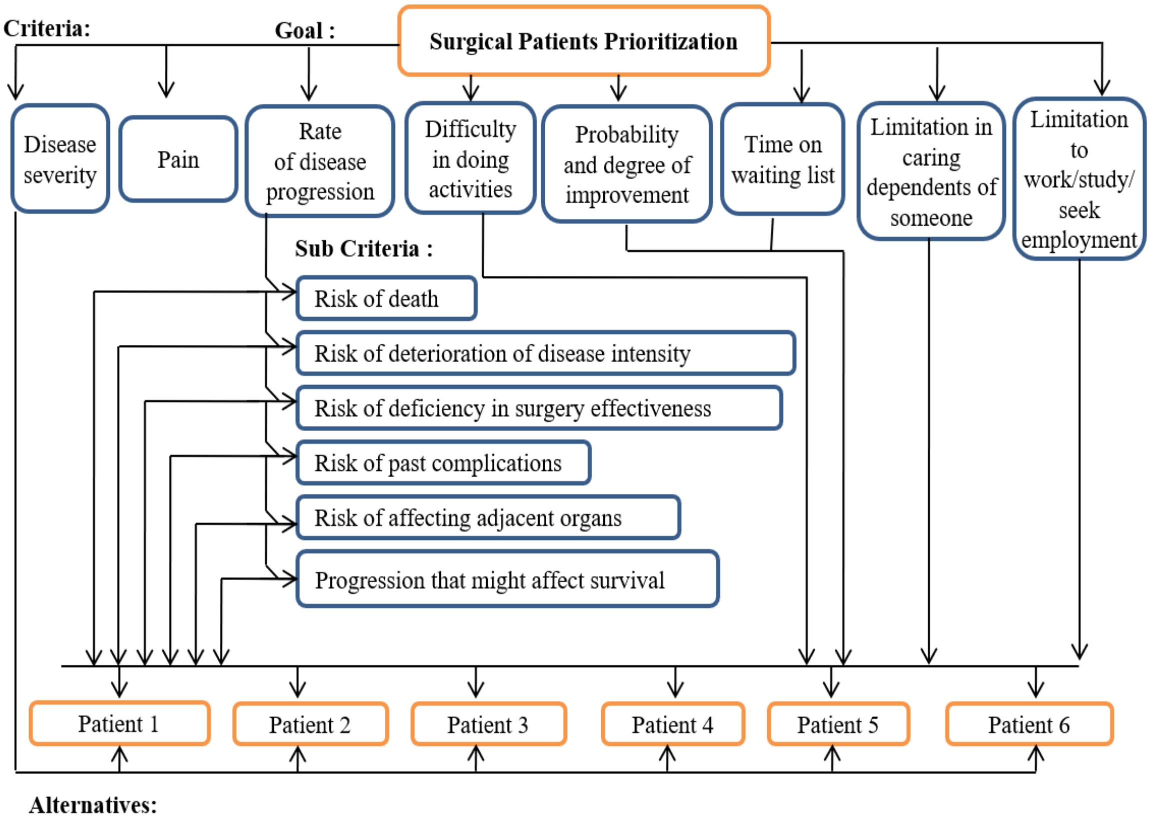


Figure 4.6: Hierarchical structure for prioritizing surgical patients

4.3.2 Phase 2 - Formation of Pair-wise Comparison Matrices (PCMs)

The selected factors are compared among themselves by experts. All experts assign individual rating to each comparison using the linguistic rating scale. Then, unanimous rating for each comparison is calculated and PCMs are designed as follows (Table 4.3).

Table 4.3: Formation of PCMs

PCM	**Size**	**Lower triangular elements of PCM**
Criteria	$(8 * 8)$	$\frac{1}{3} - \frac{1}{6} - 6 - \frac{1}{5} - 1 - \frac{1}{2} - \frac{1}{4} - \frac{1}{2} - \frac{1}{6} - \frac{1}{7} - \frac{1}{6} - \frac{1}{5} - \frac{1}{7} - \frac{1}{2} - \frac{1}{6} - \frac{1}{4} - \frac{1}{8} - \frac{1}{3} - \frac{1}{4} - \frac{1}{4} - 3 - \frac{1}{8} - \frac{1}{7} - \frac{1}{2} - \frac{1}{8} - \frac{1}{5} - 7 - 6$
Sub Criteria	$(6 * 6)$	$\frac{1}{2} - \frac{1}{7} - \frac{1}{5} - \frac{1}{8} - \frac{1}{7} - \frac{1}{3} - \frac{1}{7} - \frac{1}{6} - \frac{1}{4} - \frac{1}{3} - \frac{1}{4} - \frac{1}{5} - \frac{1}{6} - \frac{1}{2} - \frac{1}{3}$
DS	$(6 * 6)$	$\frac{1}{7} - 6 - 4 - 1 - 1 - \frac{1}{7} - \frac{1}{6} - \frac{1}{4} - \frac{1}{6} - \frac{1}{5} - \frac{1}{2} - \frac{1}{5} - \frac{1}{3} - \frac{1}{6} - 2$
Pain	$(6 * 6)$	$\frac{1}{4} - 3 - 7 - 5 - 3 - 1 - \frac{1}{6} - \frac{1}{4} - \frac{1}{5} - \frac{1}{6} - \frac{1}{2} - \frac{1}{3} - \frac{1}{4} - \frac{1}{4} - \frac{1}{5}$
RD	$(6 * 6)$	$\frac{1}{2} - 3 - 5 - 5 - 6 - \frac{1}{2} - \frac{1}{2} - \frac{1}{5} - \frac{1}{3} - \frac{1}{4} - \frac{1}{4} - \frac{1}{2} - \frac{1}{5} - \frac{1}{6} - \frac{1}{2}$
RDI	$(6 * 6)$	$3 - 4 - \frac{1}{3} - 3 - \frac{1}{2} - 1 - \frac{1}{5} - \frac{1}{6} - \frac{1}{2} - \frac{1}{6} - \frac{1}{4} - \frac{1}{7} - \frac{1}{5} - \frac{1}{5} - \frac{1}{4}$
DES	$(6 * 6)$	$3 - 1 - \frac{1}{2} - 1 - \frac{1}{3} - \frac{1}{6} - \frac{1}{2} - \frac{1}{3} - \frac{1}{4} - \frac{1}{2} - \frac{1}{2} - \frac{1}{2} - \frac{1}{2} - \frac{1}{5} - 1$
RPCs	$(6 * 6)$	$2 - 2 - \frac{1}{6} - 2 - 1 - 1 - 1 - \frac{1}{4} - \frac{1}{5} - \frac{1}{2} - 1 - \frac{1}{3} - \frac{1}{2} - \frac{1}{4} - \frac{1}{5}$
RAO	$(6 * 6)$	$4 - \frac{1}{2} - \frac{1}{4} - 1 - \frac{1}{3} - 1 - \frac{1}{4} - \frac{1}{3} - \frac{1}{5} - \frac{1}{2} - \frac{1}{3} - \frac{1}{5} - \frac{1}{3} - \frac{1}{4} - 5$
PAS	$(6 * 6)$	$\frac{1}{4} - 2 - 3 - \frac{1}{2} - 1 - \frac{1}{3} - \frac{1}{3} - \frac{1}{5} - \frac{1}{2} - \frac{1}{2} - \frac{1}{6} - \frac{1}{4} - \frac{1}{8} - \frac{1}{3} - 3$
DDA	$(6 * 6)$	$3 - 5 - 2 - 2 - 6 - \frac{1}{4} - 2 - \frac{1}{3} - \frac{1}{2} - \frac{1}{6} - 3 - 2 - 1 - 1 - 6$
PDI	$(6 * 6)$	$\frac{1}{4} - \frac{1}{7} - 2 - 1 - 5 - 2 - \frac{1}{2} - 2 - 4 - \frac{1}{4} - \frac{1}{5} - \frac{1}{6} - \frac{1}{3} - \frac{1}{3} - \frac{1}{2}$
TWL	$(6 * 6)$	$2 - \frac{1}{3} - \frac{1}{3} - 1 - \frac{1}{2} - 3 - 6 - 5 - 3 - 6 - \frac{1}{4} - \frac{1}{2} - 2 - \frac{1}{2} - \frac{1}{4}$
LCDS	$(6 * 6)$	$1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1$
LSE	$(6 * 6)$	$3 - 3 - 1 - 3 - 1 - 1 - 3 - 1 - 1 - 1 - 3 - 1 - 1 - 1 - 1$

4.3.3 Phase 3 - Consistency Measurement of Pair-wise Comparison Matrices (PCMs)

The consistency index of each PCM is measured and presented in Table 4.4. The results of Table 4.4 are presented as bar graph in Figure 4.7. Table 4.4 and bar graph in Figure 4.7 show that only two PCM, Criteria, and Disease severity (DS) fail to satisfy the CCI threshold (>=0.90). Therefore, both of the PCM's are required to be modified and optimized to satisfy the CCI threshold.

Table 4.4: CCI value of each PCM

Matrix	$\boldsymbol{C_0^*}$	$(\boldsymbol{m})$	**CCI=**$\boldsymbol{C_0^*/m}$
Criteria	6.87	8	0.8592
Sub criteria	5.73	6	0.9552
DS	5.26	6	0.8782
Pain	5.61	6	0.9352
RD	5.65	6	0.9422
RDI	5.75	6	0.9595
DES	5.51	6	0.9191
RPCs	5.48	6	0.9146
RAO	5.62	6	0.9379
PAS	5.64	6	0.9411
DDA	5.43	6	0.9053
DI	5.58	6	0.9316
TWL	5.63	6	0.9385
LCDS	6	6	1
LSE	5.94	6	0.99

4.3.4 Phase 4 - Priority Evaluation of Surgical Patients

Inconsistent PCMs are modified and combined with rest consistent PCMs to derive the priority of surgical patients are as follows:

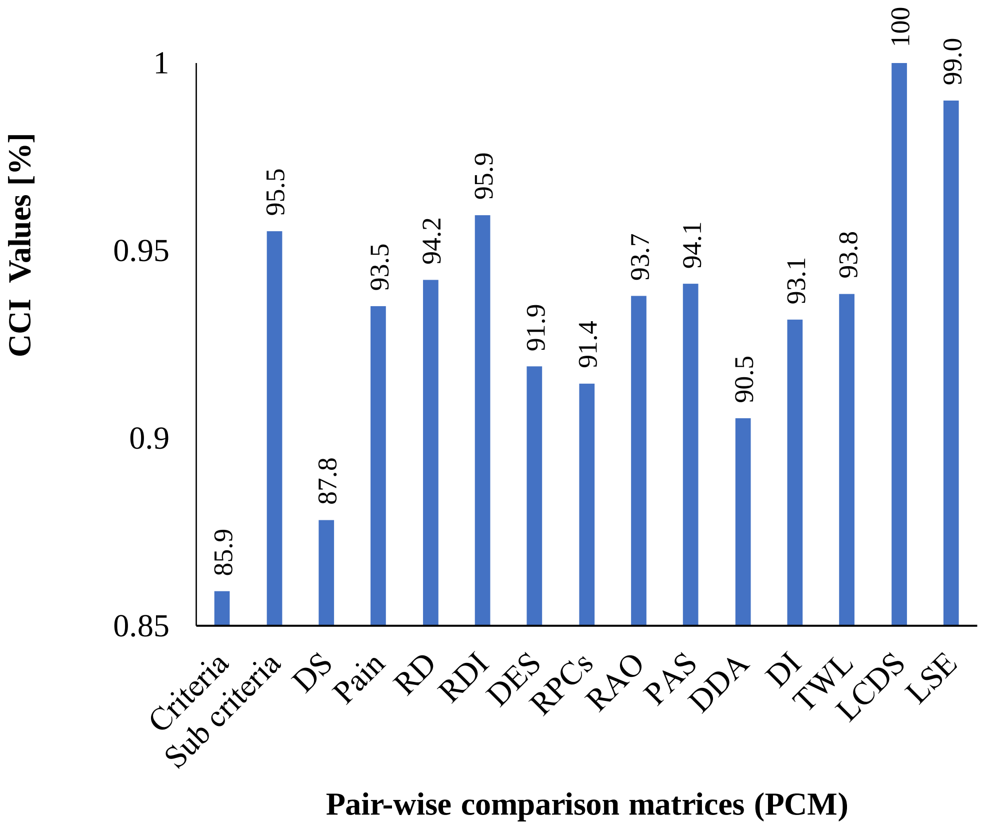


Figure 4.7: Measurement of CCI of each PCM

A. Modifications in Inconsistent PCMs

Table 4.5 shows that criteria and disease severity PCMs were initially inconsistent with CCI = 0.85 and 0.87, respectively. The proposed algorithm (HMWCA) modified these two PCMs; criteria PCM with optimized CCI = 0.908 and DI = 3.06, and disease severity PCM with optimized CCI = 0.93 and DI =2.28.

B. Evaluation of Patients' Priorities

Table 4.6 demonstrates that the proposed algorithm evaluated each surgical patient based on criteria and corresponding sub-criterion score and assigned a rank to each of them.

Table 4.5: Inconsistent PCMs and their optimized CCI

Matrix (Initial CCI)	**Optimized objective values in no of iterations**															
	10		**20**		**30**		**40**		**50**		**60**		**70**		**2-Opt**	
	CCI	**DI**	**CCI**	**DI**	**CCI**	**DI**	**CCI**	**DI**	**CCI**	**DI**	**CCI**	**DI**	**CCI**	**DI**	**CCI**	**DI**
Criteria (0.85)	0.83	17.5	0.89	19	0.9	14	0.91	10	0.9	8.3	0.90	7.4	0.90	7.3	0.90	3.06
Initial elements	$\frac{1}{3}-\frac{1}{6}-6-\frac{1}{5}-1-\frac{1}{2}-\frac{1}{4}-\frac{1}{2}-\frac{1}{6}-\frac{1}{7}-\frac{1}{6}-\frac{1}{5}-\frac{1}{7}-\frac{1}{2}-\frac{1}{6}-\frac{1}{4}-\frac{1}{8}-\frac{1}{3}-\frac{1}{4}-\frac{1}{4}-3-\frac{1}{8}-\frac{1}{7}-\frac{1}{2}-\frac{1}{8}-\frac{1}{5}-7-6$															
Optimized Elements	$\mathbf{\frac{1}{5}}-\mathbf{\frac{1}{7}}-\mathbf{3}-\mathbf{\frac{1}{2}}-1-\frac{1}{2}-\mathbf{\frac{1}{6}}-\mathbf{\frac{1}{4}}-\frac{1}{6}-\frac{1}{7}-\mathbf{\frac{1}{5}}-\mathbf{\frac{1}{3}}-\frac{1}{7}-\frac{1}{2}-\mathbf{\frac{1}{4}}-\mathbf{\frac{1}{9}}-\mathbf{\frac{1}{4}}-\mathbf{\frac{1}{6}}-\frac{1}{4}-\mathbf{\frac{1}{2}}-3-\frac{1}{7}-\mathbf{\frac{1}{7}}-\mathbf{\frac{1}{4}}-\frac{1}{8}-\frac{1}{5}-7-6$															
Disease Severity (0.87)	0.88	15.3	0.94	11.4	0.9	7.7	0.92	7	0.90	6.83	0.90	6.8	0.90	6.2	0.93	2.28
Initial Elements	$\frac{1}{7}-6-4-1-1-\frac{1}{7}-\frac{1}{6}-\frac{1}{4}-\frac{1}{6}-\frac{1}{5}-\frac{1}{2}-\frac{1}{5}-\frac{1}{3}-\frac{1}{6}-2$															
Optimized Elements	$\mathbf{\frac{1}{4}}-\mathbf{4}-\mathbf{5}-1-1-\frac{1}{7}-\mathbf{\frac{1}{5}}-\frac{1}{4}-\frac{1}{6}-\mathbf{\frac{1}{3}}-\mathbf{\frac{1}{8}}-\mathbf{\frac{1}{3}}-\mathbf{\frac{1}{5}}-\frac{1}{6}-2$															

Table 4.6: Global priority weights and final priority vector of patients

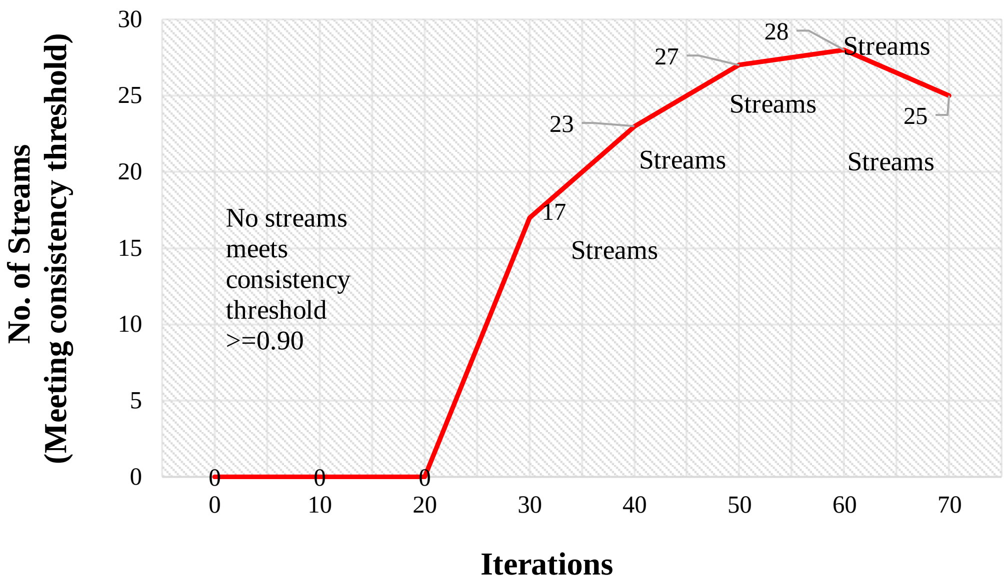
Patients	**DS**	**PAIN**	**RDP**						**DDA**	**PDI**	**TWL**	**LCDS**	**LSE**	**Final priority**	**Rank**
			RD	**RDI**	**DES**	**RPCs**	**RAO**	**PAS**							
	0.32	**0.12**	**0.07**	**0.05**	**0.03**	**0.01**	**0.01**	**0.01**	**0.15**	**0.07**	**0.03**	**0.04**	**0.08**		
P1	0.221	0.152	0.128	0.127	0.142	0.099	0.17	0.254	0.056	0.31	0.148	0.167	0.062	0.16	**6**
P2	0.158	0.103	0.116	0.358	0.301	0.298	0.381	0.155	0.127	0.106	0.159	0.167	0.188	0.161	**5**
P3	0.407	0.306	0.326	0.186	0.24	0.189	0.151	0.33	0.279	0.101	0.062	0.167	0.188	0.288	**1**
P4	0.122	0.304	0.313	0.224	0.155	0.219	0.15	0.122	0.232	0.289	0.12	0.167	0.188	0.202	**2**
P5	0.039	0.084	0.073	0.071	0.073	0.125	0.057	0.072	0.074	0.142	0.43	0.167	0.188	0.092	**4**
P6	0.053	0.051	0.045	0.035	0.088	0.069	0.092	0.067	0.232	0.053	0.081	0.167	0.188	0.096	**3**

4.3.5 Performance Evaluation of Proposed Prioritization Model

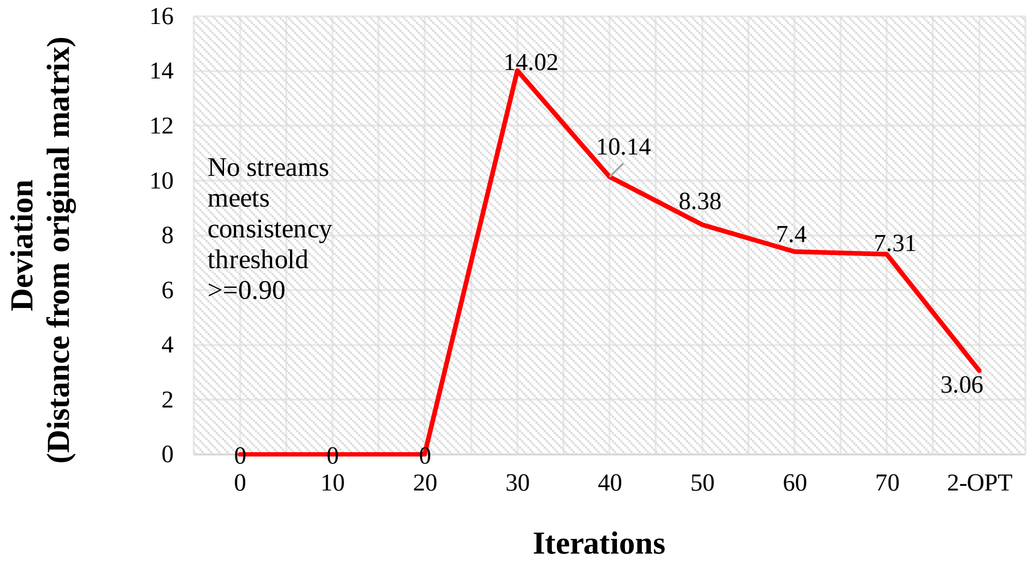
The performance of the proposed prioritization model is evaluated in two steps: Firstly, the performance of the proposed hybrid algorithm (HMWCA) is evaluated, and Secondly, performance of proposed prioritization model is measured. Since the indicator of an optimal solution is getting both the smallest deviation (DI) and satisfying the CCI threshold. The performance of the proposed algorithm is evaluated using three criteria: Improvement in CCI, Optimal CCI score, and paired sample t-test. Figure 4.8 (a) (Criteria PCM) and Figure 4.8 (c) (Disease Severity PCM) depicts that in the beginning iterations, no genotype (streams representing solution) could achieve the CCI threshold (>=0.90). After the 20th iteration, the genotypes start satisfying the CCI threshold. Although in some iterations, there may have been circumstances where the number of genotypes satisfying the threshold decreased. Nevertheless, the genotypes obtained a consistent position. After the genotypes achieve the CCI threshold, the next focus is to get the minimal deviation (DI) and minimal deviation score $F(x)$ as well. Figure 4.8 (b) (Criteria PCM) and Figure 4.8 (d) (Disease Severity PCM) depicts that the DI obtained from the 30th iteration until the 2-opt iterations. The DI of iteration 1 to 20 is ignored to determine the $F(x)$ as there are no genotypes that could satisfy the CCI threshold. These figures show some iterations, where $F(x)$ is minimal. During each iteration, $F(x)$ gradually decreases. Figures 4.9 (a) and (b) shows the best optimal position, achieved after 2-opt iterations, with $F(x)$ = 3.58 and $F\ (x)$ = 2.79, respectively. Subsequently, the performance of our model, PSWL-CCI, is evaluated using two criteria: ED and MV. The performance of our method is then compared (Table 4.9) with those of other prioritization methods: EV, WLS, LLS, and AN. The inputs are; PCM $Z = \left(Z_{pq}\right)_{k*m}$ and the priority vector $W\ = (\overline{w_1}, \overline{w_2,}, \dots \overline{w_m})^T$. The outputs are the values of the error measures (ED and MV) between PCM (Z) and derived priority vector $W\ = (\overline{w_1}, \overline{w_2,}, \dots \overline{w_m})^T$. Thus, the prioritization method with a lower error value is considered superior to the others. The results of performance comparison of Table 4.9 are presented as bar graph in Figure

4.10, which evidence that the proposed PSWL-CCI model has the lowest values of the error measures: ED and MV among all the other prioritization methods studied. The results validate that the proposed PSWL-CCI model is better than the existing methods. Therefore, the resulting priority of a patient for surgery, from Table 4.6 is as follows:

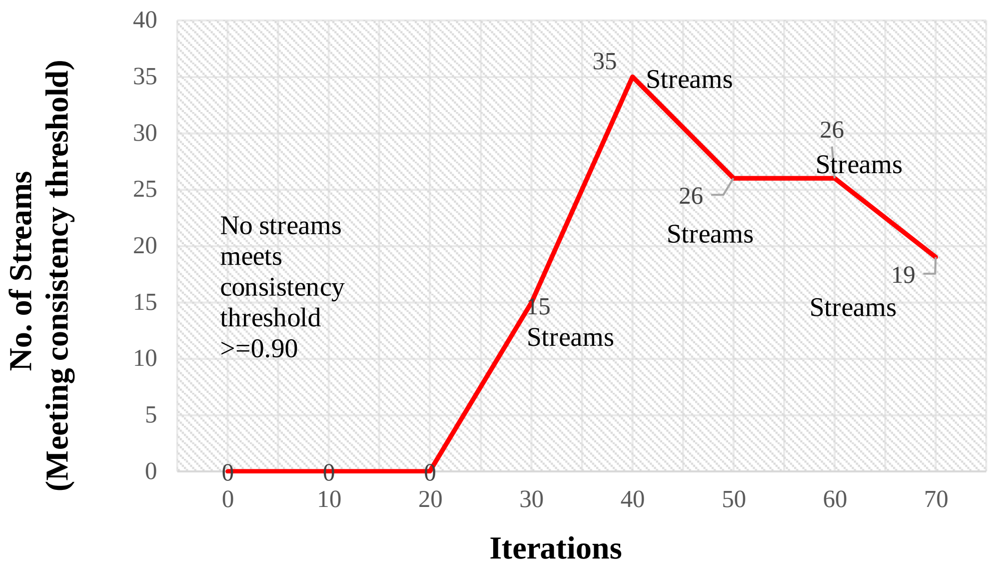
P3 > P4 > P6 > P5 > P2 > P1



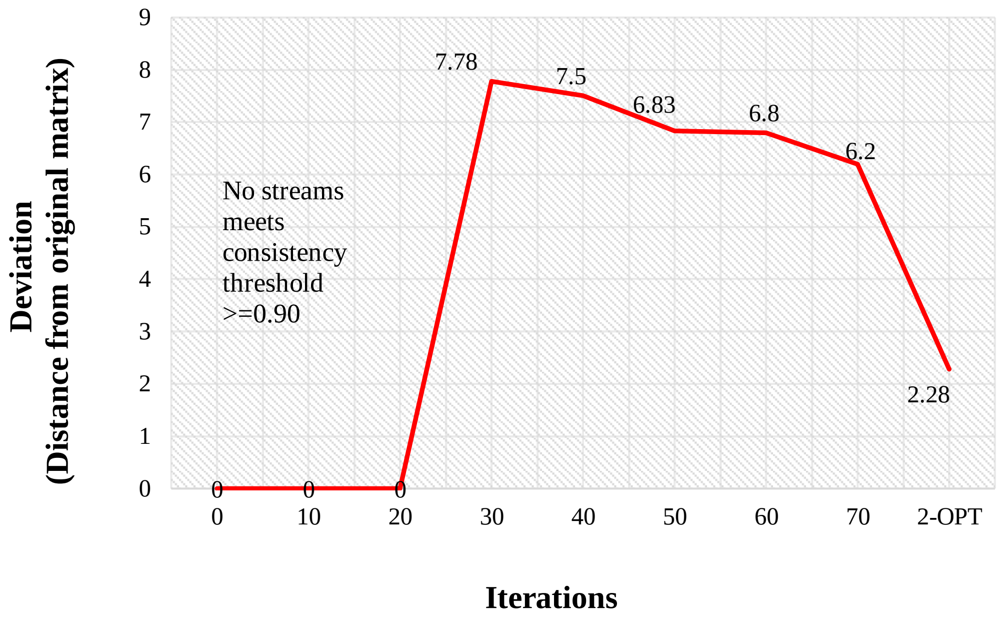
(a) Criteria PCM: No. of solutions satisfying CCI threshold



(b) Criteria PCM: No. of solutions with deviation



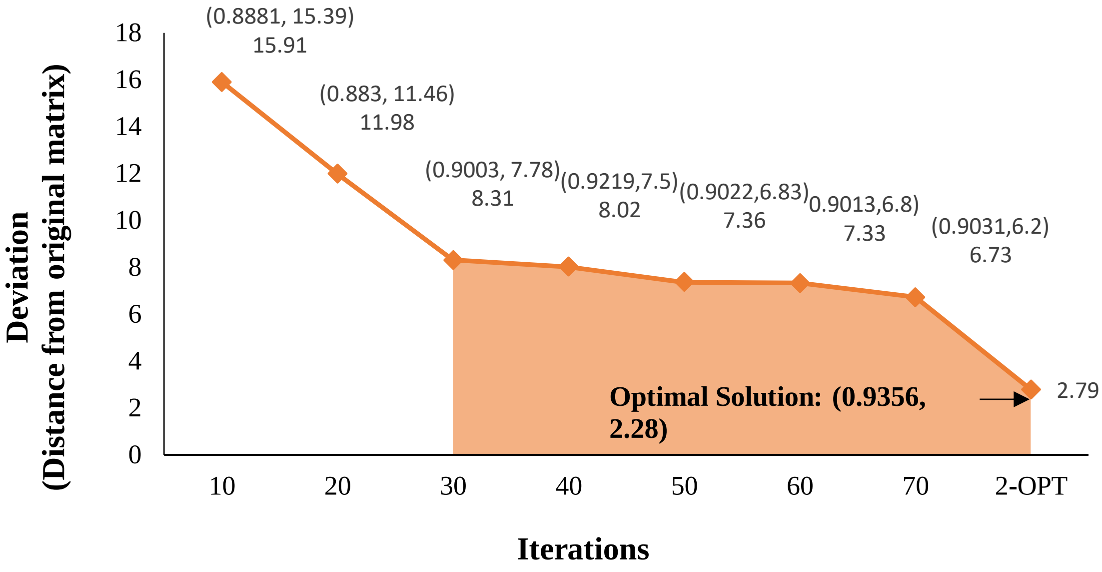
(c) Disease Severity PCM: No. of solutions satisfying CCI threshold



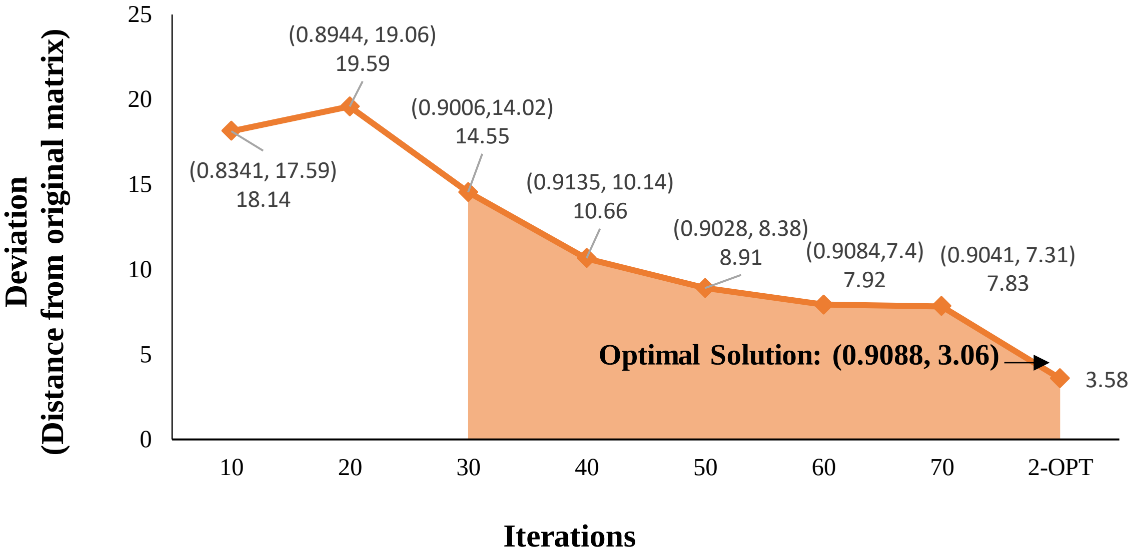
(d) Disease Severity PCM: No. of solutions with deviation

Figure 4.8: Performance of proposed hybrid algorithm on patients' dataset

(Improvement in CCI)



(a) Criteria PCM: Smallest deviation satisfying CCI threshold



(b) Disease severity PCM: Smallest deviation satisfying CCI threshold

Figure 4.9: Performance of proposed hybrid algorithm (Optimal CCI score)

Table 4.7: Performance comparison of the proposed hybrid algorithm with existing algorithms

Matrix **[Reference]** **(size)**	**Initial element** **(Initial CCI)**	**Optimized CCI**			**Improvement in CCI** **(Proposed Algorithm)**
		CCI Optimized Elements (CCI, DI, Deviation score, Iterations)	**WCA Optimized Elements (CCI, DI, Deviation Score, Iterations)**	**Proposed HMWCA Optimized Elements (CCI, DI, Deviation Score, Iterations)**	
PCM2[90] $(\mathbf{7} * \mathbf{7})$	$\frac{1}{5}-7-7-\frac{1}{5}-\frac{1}{5}-\frac{1}{9}$ $-\frac{1}{7}-\frac{1}{3}-9-\frac{1}{9}-5$ $-9-\frac{1}{7}-\frac{1}{7}-\frac{1}{9}-7$ $-\frac{1}{3}-7-9-5-9$	$\frac{1}{5}-7-7-\frac{1}{5}-\frac{1}{5}-\frac{1}{9}$ $-\frac{1}{7}-\frac{1}{3}-\mathbf{\frac{1}{8}}-\mathbf{2}-\mathbf{\frac{1}{3}}-\mathbf{\frac{1}{7}}$ $-\frac{1}{7}-\frac{1}{7}-\frac{1}{9}-\mathbf{5\frac{3}{4}}-\frac{1}{3}$ $-\mathbf{\frac{2}{9}}-9-5-9$	$\mathbf{\frac{1}{2}}-7-\mathbf{6}-\mathbf{\frac{1}{3}}-\mathbf{\frac{1}{6}}-\mathbf{\frac{1}{5}}$ $-\mathbf{\frac{1}{2}}-\frac{1}{3}-\mathbf{\frac{1}{3}}-\mathbf{\frac{1}{2}}-\mathbf{\frac{1}{4}}$ $-\mathbf{\frac{1}{3}}-\frac{1}{7}-\mathbf{\frac{1}{8}}-\mathbf{\frac{1}{7}}-\mathbf{6}$ $-\mathbf{\frac{1}{2}}-\mathbf{\frac{1}{2}}-\mathbf{6}-\mathbf{4}-\mathbf{8}$	$\mathbf{\frac{1}{2}}-7-6-\mathbf{\frac{1}{3}}-\mathbf{\frac{1}{6}}-\frac{1}{9}$ $-\mathbf{\frac{1}{5}}-\mathbf{\frac{1}{3}}-\frac{1}{3}-\mathbf{\frac{1}{2}}-\mathbf{\frac{1}{2}}$ $-\mathbf{\frac{1}{6}}-\frac{1}{7}-\mathbf{\frac{1}{3}}-\mathbf{\frac{1}{6}}-\mathbf{6}$ $-\mathbf{\frac{1}{2}}-\mathbf{\frac{1}{2}}-9-5-9$	
	(0.7283)	**(0.9023, 15.16, 15.68, 40)**	**(0.9101, 15.12, 15.6, 40)**	**(0.9156, 14.76, 15.28, 40)**	**0.1873**
PCM11[91] $(\mathbf{5} * \mathbf{5})$	$3-\frac{1}{2}-\frac{1}{7}-6-9-9$ $-2-4-4-5$	$\mathbf{2}-\mathbf{\frac{1}{8}}-\frac{1}{7}-\mathbf{\frac{1}{6}}-\mathbf{\frac{1}{2}}-\mathbf{6}-$ $\mathbf{\frac{1}{4}}-\mathbf{\frac{1}{2}}-\mathbf{7\frac{5}{7}}-\mathbf{6\frac{1}{2}}$	$3-\frac{1}{2}-\frac{1}{7}-\mathbf{4}-\mathbf{5}-\mathbf{6}$ $-\mathbf{5}-\mathbf{3}-\mathbf{7}-\mathbf{4}$	$3-\frac{1}{2}-\frac{1}{7}-6-9-9$ $-\mathbf{4}-\mathbf{5}-\mathbf{6}-5$	
	(0.8616)	**(0.9166, 12.15, 12.67,10)**	**(0.9364, 7, 7.51, 10)**	**(0.9015, 3, 3.52, 10)**	**0.0399**
PCM14[98] $(\mathbf{6} * \mathbf{6})$	$\frac{1}{5}-1-5-3-7-1$ $-7-3-7-3-3$ $-5-\frac{1}{5}-\frac{1}{5}-5$	$\mathbf{\frac{2}{5}}-\mathbf{\frac{1}{2}}-\mathbf{\frac{1}{2}}-\mathbf{6}-\mathbf{\frac{3}{8}}-\mathbf{4}$ $-\mathbf{5}-\mathbf{4}-\mathbf{5}-\mathbf{7}-\mathbf{4}-\mathbf{2}$ $-\mathbf{\frac{1}{2}}-\mathbf{\frac{1}{4}}-\mathbf{\frac{1}{6}}$	$\mathbf{\frac{1}{2}}-\mathbf{6}-\mathbf{4}-\mathbf{5}-\mathbf{6}-\mathbf{2}$ $-\mathbf{8}-\mathbf{8}-\mathbf{2}-\mathbf{3}-\mathbf{4}$ $-\mathbf{5}-\mathbf{\frac{1}{2}}-\mathbf{\frac{1}{3}}-\mathbf{2}$	$\mathbf{\frac{1}{2}}-\mathbf{3}-\mathbf{4}-\mathbf{5}-\mathbf{6}-\mathbf{4}$ $-7-\mathbf{6}-\mathbf{5}-\mathbf{2}-\mathbf{4}$ $-5-\mathbf{\frac{1}{2}}-\mathbf{\frac{1}{3}}-\mathbf{2}$	
	(0.7934)	**(0.9233, 11.88, 12.4, 42)**	**(0.9009, 9.65, 10.17, 42)**	**(0.9026, 6.57, 7.09, 42)**	**0.1092**

PCM15[98] $(\mathbf{6} * \mathbf{6})$	$\frac{1}{5}-\frac{1}{9}-3-1-5-\frac{1}{5}-5-5-5-7-3-3-\frac{1}{3}-3-7$ **(0.7767)**	$\boldsymbol{\frac{1}{3}-\frac{1}{2}-4-5-\frac{1}{2}-\frac{1}{4}-\frac{1}{2}-7-\frac{1}{3}-8-8-9-9-7-6}$ **(0.9256, 15.41, 15.9, 30)**	$\frac{1}{5}-\mathbf{\frac{1}{2}}-\mathbf{4}-\mathbf{\frac{1}{4}}-\mathbf{3}-\mathbf{\frac{1}{6}}-\mathbf{3}-\mathbf{6}-\mathbf{3}-\mathbf{9}-\mathbf{2}-\mathbf{8}-\mathbf{\frac{1}{2}}-\mathbf{7}-\mathbf{2}$ **(0.9092, 9.26, 9.78, 30)**	$\boldsymbol{\frac{1}{7}-\frac{1}{2}-4-\frac{1}{5}-3-\frac{1}{3}-2-7-3-6-1-2-\frac{1}{2}-3-2}$ **(0.9122, 7.34, 7.86, 30)**	**0.1355**
Criteria $(\mathbf{6} * \mathbf{6})$	$\frac{1}{3}-\frac{1}{6}-6-\frac{1}{5}-1-\frac{1}{2}-\frac{1}{4}-\frac{1}{2}-\frac{1}{6}-\frac{1}{7}-\frac{1}{6}-\frac{1}{5}-\frac{1}{7}-\frac{1}{2}-\frac{1}{6}-\frac{1}{4}-\frac{1}{8}-\frac{1}{3}-\frac{1}{4}-\frac{1}{4}-3-\frac{1}{8}-\frac{1}{7}-\frac{1}{2}-\frac{1}{8}-\frac{1}{5}-7-6$ **(0.85)**	$\mathbf{2}-\mathbf{3}\,\mathbf{\frac{1}{3}}-\mathbf{3}-\mathbf{\frac{1}{3}}-1-\mathbf{\frac{1}{6}}-\mathbf{\frac{1}{6}}-\mathbf{\frac{1}{7}}-\frac{1}{6}-\frac{1}{7}-\mathbf{\frac{1}{5}}-\mathbf{\frac{1}{2}}-\mathbf{\frac{1}{8}}-\mathbf{\frac{1}{4}}-\mathbf{\frac{1}{4}}-\mathbf{\frac{1}{5}}-\mathbf{\frac{1}{5}}-\mathbf{\frac{1}{5}}-\mathbf{2}\,\mathbf{\frac{2}{9}}-\mathbf{3}\,\mathbf{\frac{1}{3}}-\mathbf{4}-\mathbf{\frac{1}{6}}-\frac{1}{7}-\mathbf{\frac{1}{4}}-\mathbf{2}-\frac{1}{5}-\mathbf{\frac{1}{2}}-\mathbf{\frac{1}{2}}$ **(0.9216, 10.62, 11.14, 30)**	$\mathbf{\frac{1}{5}}-\mathbf{\frac{1}{6}}-\mathbf{3}-\mathbf{\frac{1}{2}}-1-\frac{1}{2}-\mathbf{\frac{1}{6}}-\mathbf{\frac{1}{4}}-\frac{1}{6}-\frac{1}{7}-\mathbf{\frac{1}{5}}-\mathbf{\frac{1}{2}}-\mathbf{\frac{1}{5}}-\frac{1}{2}-\mathbf{\frac{1}{4}}-\mathbf{\frac{1}{6}}-\mathbf{\frac{1}{5}}-\mathbf{\frac{1}{6}}-\frac{1}{4}-\mathbf{\frac{1}{2}}-\mathbf{4}-\frac{1}{8}-\frac{1}{7}-\mathbf{\frac{1}{4}}-\mathbf{\frac{1}{3}}-\frac{1}{5}-7-\mathbf{7}$ **(0.9013, 3.39, 3.91, 30)**	$\mathbf{\frac{1}{5}}-\mathbf{\frac{1}{7}}-\mathbf{3}-\mathbf{\frac{1}{2}}-1-\frac{1}{2}-\mathbf{\frac{1}{6}}-\mathbf{\frac{1}{4}}-\frac{1}{6}-\frac{1}{7}-\mathbf{\frac{1}{5}}-\mathbf{\frac{1}{3}}-\frac{1}{7}-\frac{1}{2}-\frac{1}{4}-\mathbf{\frac{1}{9}}-\frac{1}{4}-\mathbf{\frac{1}{6}}-\frac{1}{4}-\mathbf{\frac{1}{2}}-3-\mathbf{\frac{1}{7}}-\mathbf{\frac{1}{7}}-\mathbf{\frac{1}{4}}-\frac{1}{8}-\frac{1}{5}-7-6$ **(.908, 3.06, 3.58, 30)**	**0.052**
DS $(\mathbf{6} * \mathbf{6})$	$\frac{1}{7}-6-4-1-1-\frac{1}{7}-\frac{1}{6}-\frac{1}{4}-\frac{1}{6}-\frac{1}{5}-\frac{1}{2}-\frac{1}{5}-\frac{1}{3}-\frac{1}{6}-2$ **(0.87)**	$\mathbf{\frac{1}{5}}-\mathbf{6}\,\mathbf{\frac{1}{2}}-\mathbf{6}-1-\mathbf{3}-\mathbf{\frac{1}{5}}-\mathbf{2}-\mathbf{5}-\mathbf{\frac{1}{6}}-\frac{1}{5}-\mathbf{2}-\mathbf{6}-\mathbf{\frac{1}{7}}-\mathbf{\frac{2}{3}}-\mathbf{\frac{1}{3}}$ **(0.9369, 8.56, 9.07, 30)**	$\mathbf{\frac{1}{3}}-\mathbf{2}-\mathbf{5}-1-1-\mathbf{\frac{1}{4}}-\mathbf{\frac{1}{2}}-\frac{1}{4}-\mathbf{\frac{1}{3}}-\mathbf{\frac{1}{3}}-\mathbf{\frac{1}{5}}-\mathbf{\frac{1}{3}}-\mathbf{\frac{1}{5}}-\frac{1}{6}-2$ **(0.9451, 4.16, 4.67, 30)**	$\mathbf{\frac{1}{4}}-\mathbf{4}-\mathbf{5}-1-1-\frac{1}{7}-\mathbf{\frac{1}{5}}-\frac{1}{4}-\frac{1}{6}-\mathbf{\frac{1}{3}}-\mathbf{\frac{1}{8}}-\mathbf{\frac{1}{3}}-\mathbf{\frac{1}{5}}-\frac{1}{6}-2$ **(0.9356, 2.28, 2.79, 30)**	**0.06**

Table 4.7 comprehends that the proposed algorithm (HMWCA), implemented on different inconsistent PCM, produces more optimal results as compared to existing CCI improvement methods. Results obtained in Table 4.8 show that for each comparison of the proposed algorithm (HMWCA), P-value is less than 0.05. It validates that there is a significant improvement in the CCI threshold by implementing the proposed hybrid algorithm.

Table 4.8: Paired-Sample T-Test

Paired values (Sample size)	**Description**	**Paired Difference**					**T**	**DF**	**P-value**
		Mean	**Std. Deviation**	**Std. Error mean**	**95% Confidence Interval of difference**				
					Lower	**Upper**			
Optimized – Initial Values (6)	CCI (Original values- HMWCA)	0.096	0.057	0.023	0.035	0.15	4.0	5	0.009
	DI (Original values HMWCA)	6.193	4.669	1.906	1.292	11.093	3.2	5	0.02
	Optimized CCI score (Original values-HMWCA)	6.713	4.669	1.906	1.812	11.613	3.52	5	0.01
	Optimized CCI score (CCI (WAM)-HMWCA)	6.09	3.11	1.27	2.83	9.36	4.8	5	0.004
	Optimized CCI score (ER-WCA- HMWCA)	2.05	1.48	0.606	0.498	3.617	3.39	5	0.01

Table 4.9: Performance of proposed PSWL-CCI model in comparison to existing prioritization methods

Performance Measure	**PSWL-CCI (Proposed model)**	**EV Method**	**WLS Method**	**AN Method**	**LLS Method**
ED	4.6449 **(1)**	5.5179 **(4)**	7.7789 **(5)**	4.8669 **(2)**	5.3119 **(3)**
MV	1	2	2	1	2

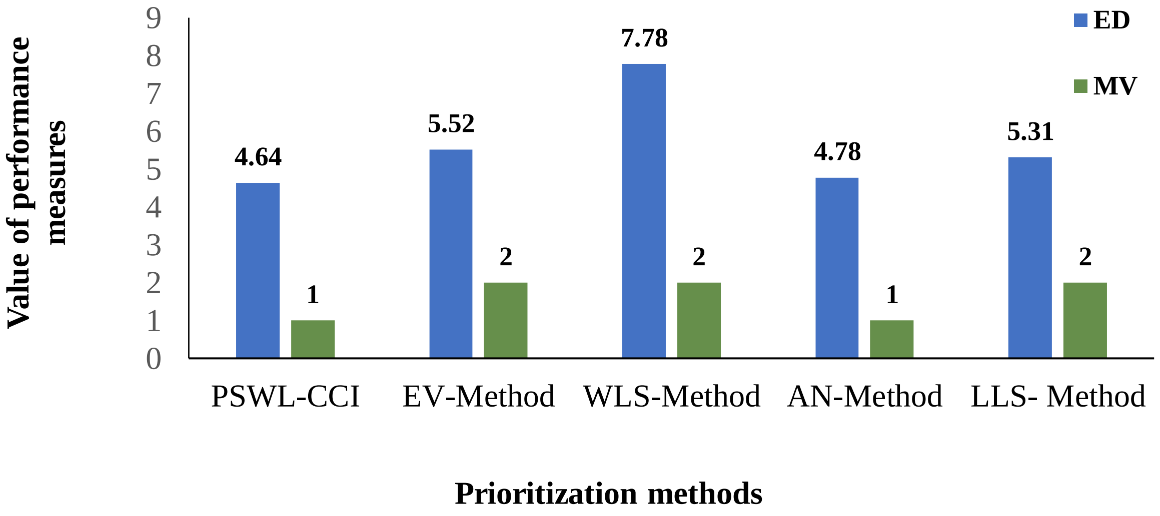


Figure 4.10: Performance comparison of PSWL-CCI model with other existing prioritization methods

Chapter 5

CLUSTER ANALYSIS OF SURGICAL PATIENTS

Focusing on finding the similarities among the surgical patients that can assist healthcare experts in surgical decision-making, we have proposed an automatic clustering algorithm. The proposed algorithm is based on a population-based metaheuristic algorithm called the artificial electric field algorithm (AEFA) and deals with mixed datasets. The proposed clustering algorithm aims to find the optimal number of cluster partitions automatically. It utilizes a real-coded variable-length candidate solution to detect the optimal number of clusters automatically. The concepts of threshold setting and cut-off ratio are used in the optimization process to refine the clusters. The similarity between data points and different cluster centers is measured using Euclidean distance (for numeric attributes) and the probability of co-occurrence of values (for categorical attributes).

5.1 Clustering Analysis and Optimization Methods

This section briefly discusses the basic concepts of partitioning clustering, distance measure for the mixed dataset, and artificial electric field algorithm (AEFA) for optimization.

5.1.1 Partitioning Clustering

Partitioning clustering is a data clustering approach that groups the data points into disjoint clusters. Let us consider a dataset $Z = \{Z_1, Z_2, Z_3, \dots, Z_n\}$ of n data points. In the dataset, each data point is represented by D attributes. For example, $Z_j = (Z_{j1}, Z_{j2}, \dots, Z_{jD})$ is a vector representing the j^{th}

data point, and Z_{ji} represents the i^{th} attribute of Z_j. The objective of the partitional clustering algorithm is to determine the disjoint cluster (C_i) that satisfies the following condition:

$$C_i \neq \phi, i = 1,2 \dots k, \tag{5.1}$$

$$\sum_{i=1}^{k} C_i = Z, \tag{5.2}$$

$$C_i \cap C_j \neq \phi \ \forall i, j \tag{5.3}$$

Where, k represents the number of clusters.

5.1.2 Distance Measure for Mixed Datasets

The belonging of a datapoint with a cluster is measured using the distance computed between the cluster and the data point. Distance measure ensures the similarity among the data points belonging to the same cluster and dissimilarity between disjoint clusters. Grouping a mixed dataset into clusters is a significantly challenging task. Since it requires a suitable distance measure to effectively compute the similarity/dissimilarity between data points and cluster centers, this chapter utilizes the distance measure proposed by Ahmad and Dey[121] in the clustering process. The distance (ϑ) between any data point (d_i) and cluster center (c_j) is computed as follows:

$$\vartheta(d_i, c_j) = \sum_{t=1}^{m_r} \left(w_t \left(d_{it}^r - c_{jt}^r \right) \right)^2 + \sum_{t=1}^{m_c} \Omega\left(d_{it}^c, C_{jt}^c \right)^2 \tag{5.4}$$

Here, the first component represents the distance between t^{th} numeric attribute value of data point d_i; (d_{it}^r) and cluster center (c_{jt}^r) and the second component represents the distance between the t^{th} categorical attribute value of data point d_i; (d_{it}^c) and cluster center (c_{jt}^c). w_t represents the significance of t^{th} numeric attribute. It is calculated by computing the average distance between all pairs of discretized numeric attribute values. The distance between two attribute values of a categorical attribute is measured by computing the co-occurrence of these values with attribute values of other categorical attributes.

5.1.3 Artificial Electric Field Algorithm (AEFA)

AEFA[146] is a population-based meta-heuristic, which mimics Coulomb's law of electrostatic attraction force and the law of motion. In AEFA, the possible candidate solutions of the given problem are represented as a collection of the charged particles. The charge associated with each charged particle helps in determining the performance of each candidate solution. Attraction electrostatic force causes each particle to attract towards one another resulting in the global movement towards particle with the heavier charge. A candidate solution to the problem represents the position of charged particles, where the charge of particles is determined using a fitness function. The steps of AEFA are as follows:

A. Initialization of Population

A population of P candidate solutions (charged particle) is initialized as follows:

$$CP_i = \left(CP_i^1, CP_i^2, CP_i^k \ldots \ldots \ldots CP_i^d\right), \forall\, i = 1,2,3 \ldots p \tag{5.5}$$

Where, CP_i^k represents the position of i^{th} charged particle in the k^{th} dimension, and D is the dimensional space.

B. Fitness Evaluation

The performance of each charged particle depends on the fitness values at each iteration. The best and the worst fitness is computed as follows:

$$Best\ (t) = \ min_{i=1}^{n}\ Fitness\ (t) \tag{5.6}$$

$$Worst\ (t) = \ max_{i=1}^{n}\ Fitness\ (t) \tag{5.7}$$

Where, $Fitness\ (t)$ and $n$are the fitness value of i^{th} charged particle and size of the population, respectively at time t. $Best\ (t)$, $Worst\ (t)$ represents the best fitness and the worst fitness of all charged particles at time t.

C. Computation of Coulomb's Constant

At time t, the Coulomb's constant is denoted by $K_c(t)$ and computed as follows:

$$K_c(t) = \ K_0 * \exp\left(-\alpha \frac{iter}{N}\right) \tag{5.8}$$

Where, K_0 represents initial value and α is a random value, respectively. $iter$ and N represents current iteration and maximum number of iterations.

D. Computation of the Charge of Charged Particles

At time T, the charge of i^{th} charged particle is represented by $Q_i(T)$. It is computed based on the current population's fitness as follows:

$$Q_i(T) = \frac{q_i(t)}{\sum_{i=1}^{n} q_i(t)} \tag{5.9}$$

Where, $$q_i(t) = \exp\left(\frac{fitness_{cp_i}(t) - Worst(t)}{Best(t) - Worst(t)}\right) \tag{5.10}$$

E. Computation of the Electrostatic Force and Acceleration of the Charged Particles

1. The electrostatic force exerted by j^{th} charged particle on the i^{th} charged particle in the D^{th}

dimension at time T is computed as:

$$F_{ij}^{D}(t) = K(t)\frac{\left(Q_i(t) * Q_j(t)\right) * \left(P_j^{D}(t) - X_j^{D}(t)\right)}{R_{ij}(t) + \varepsilon} \tag{5.11}$$

$$F_i^{D}(t) = \sum_{j=1, j\neq i}^{N} rand\,() * F_{ij}^{D}(t) \tag{5.12}$$

Where, $Q_i(t)$ and $Q_j(t)$ are the charges of i^{th} and j^{th} charged particle at any time t. ε is a small positive constant and $R_{ij}(t)$ is the distance between two charged particles i and j. $P_j^{D}(t)$ and $X_j^{D}(t)$ are the global best and current position of the charged particle at time t. $F_i^{D}(t)$ is the net force exerted on i^{th} charged particle by all other charged particles at time t. $rand()$ is a uniform random number in [0, 1] interval.

2. The acceleration $a_i^{D}(t)$ of i^{th} charged particle at time T in D^{th} dimension is computed using the Newton law of motion as follows:

$$a_i^{D}(t) = \frac{Q_i(T) * E_i^{D}(t)}{M_i^{D}(t)}, \quad E_i^{D}(t) = \frac{F_i^{D}(t)}{Q_i(T)} \tag{5.13}$$

Where, $E_i^{D}(t)$ and $M_i^{D}(t)$ represents the electric field and unit mass of i^{th} charged particle at any time and in D^{th} dimension respectively.

F. Updation of Velocity and Position of Charged Particle

At time T, the position and velocity of i^{th} charged particle in D^{th} dimension is updated as follows:

$$vel_i^{D}(T+1) = rand_i * vel_i^{D}(T) + a_i^{D}(t) \tag{5.14}$$

$$CP_i^{D}(T+1) = CP_i^{D}(T) + vel_i^{D}(t) \tag{5.15}$$

where, $rand()$ is a uniform random number in the interval [0, 1].

5.2 Proposed Clustering Algorithm

In this section, the proposed clustering algorithm is described in detail. The algorithm starts with accepting preprocessed mixed dataset as input and initializes parameters. Subsequently, a population of candidate solutions is generated, where each solution is composed of two segments: the first segment represents threshold values, and the second segment represents cluster centers. The threshold values of the first segment determine whether or not the corresponding cluster centers are active in the second segment. Further, by computing the fitness value for each candidate solution, the best solution is selected. The population is iteratively updated until the termination condition is satisfied and the optimal solution is returned. The symbols used in the proposed clustering algorithm are presented in Table 5.1. The detailed workflow and pseudocode of the proposed clustering algorithm are presented in Figure 5.1 and Algorithm 6, respectively.

5.2.1 Population Generation

In this step, the population of candidate solutions is initialized. For the dataset which has n data point with dimension D each, the maximum number of cluster center is computed as follows:

$$K_{max} = \sqrt{n} \tag{5.16}$$

Each candidate solution (X_i) is represented by $D_x = K_{max} + K_{max} * D$ dimension. The first segment of X_i contains K_{max} threshold values (Th), which are generated randomly within the interval of [0,1]. The second segment comprises the k_{max} cluster centers, where each of which has D dimension. The Threshold values representation of each candidate solution (X_i) is as follows:

$$\underbrace{\begin{array}{|c|c|c|c|}\hline Th_{i1} & Th_{ij} & .. & Th_{ik_{max}} \\ \hline \end{array}}_{\text{Threshold values}}\underbrace{\begin{array}{|c|c|c|c|c|}\hline C_{i1} & C_{i2} & C_{ij} & .. & C_{ik_{max}} \\ \hline \end{array}}_{\text{Cluster centers}}$$

Where, Th_{ij} represent the threshold value of C_{ij} cluster center and C_{ij} represents j^{th} cluster center of i^{th} candidate solution.

Table 5.1: Symbols used in the proposed clustering algorithm

Symbol	**Definition**
ϑ	Distance between the data point and cluster center
d_{it}^{r}	t^{th} numeric value of i^{th} datapoint
d_{it}^{c}	t^{th} categorical value of i^{th} datapoint
CP_i^D	Position of i^{th} charged particle (candidate solution) in D^{th} dimension
CP_i^J	j^{th} cluster center belonging to i^{th} charged particle
D	Dimension of objective space that helps in determining the overall dimension of a candidate solution
Fitness()	Objective (fitness) function
$Best\ (t)$	The best fitness value of a charged particle at time t
$Worst\ (t)$	The worst fitness value of a charged particle at time t
K_0	Initial value of Coulomb’s constant
K_c	Coulomb’s constant at time t
$q_i(t)$	Small value of charge on i^{th} charged particle at time t that helps in determining the total charge of (i^{th}) charged particle
$Q_i(t)$	Total charge acting on a i^{th} charged particle at time t
$F_{ij}^{DS}(t)$	Force exerted by j^{th} charge particle on i^{th} charge particle in D^{th} dimension at time t
$F_i^D(t)$	Net force on i^{th} charged particle in D^{th} dimension at time t
$P_i^D(t)$	Global best position of j^{th} charged particle in D^{th} dimension at time t
$X_J^D(t)$	Current position of j^{th} charged particle in D^{th} dimension at time t
$R_{ij}(t)$	Distance between i^{th} and j^{th} charged particle at time t
k_{active}	Total number of active clusters
$a_i^D(t)$	Acceleration of i^{th} charged particle in D^{th} dimension at time t

$E_i^D(t)$	Electric field of i^{th} charged particle in D^{th} dimension at time *t*
$M_i^D(t)$	Mass of i^{th} charged particle in D^{th} dimension at time *t*
$vel_i^D(t)$	Velocity of i^{th} charged particle in D^{th} dimension at time *t*
K_{max}	Maximum number of clusters centers
n	Population size (number of datapoints)
D_x	Overall dimension of a charged particle (candidate solution)
N	Maximum number of iterations
Th_i^j	Threshold value of j^{th} cluster center belonging to i^{th} candidate solution
Th_{cl}	The selection threshold value of a cluster center
T_{cot}	The cut-off threshold value of a cluster center
$SD(CP_i)$	SD index value of a i^{th} charged particle
$Scat(CP_i)$	Intra-cluster distance of a i^{th} charged particle
$Dist(CP_i)$	Inter-cluster distance a i^{th} charged particle
$\sigma_{CP_i}^D$	Variance of clusters belonging to i^{th} charged particle
σ_x^D	Variance of dataset (x)
D_{max}	Maximum distance between the cluster centers of a charged particle
D_{min}	Minimum distance between the cluster centers of a charged particle

5.2.2 Active Cluster Center Selection

In this step, active cluster centers among k_{max} cluster centers of a candidate solution are selected. Selection of the active cluster centers is based on the following condition:

$$C_{ij} = \begin{cases} 1\,, & Th_i^j > T_{cot} \\ 0, & Otherwise \end{cases} \tag{5.17}$$

Where, T_{cot} represents the cutoff threshold for every cluster center. It is initially set to a random value between [0,1] interval. In succeeding iterations T_{cot} is computed as follows:

$$T_{cot} = \frac{1}{k} \sum_{l=1}^{K_{active}} Th_{cl} \tag{5.18}$$

$$Th_{cl} = \sqrt{\frac{1}{n_{cl}}\left((\sum_{i=1}^{m_r} \Omega(d_i^r, CP_{cl}^r)^2) + \left(\sum_{i=1}^{m_c} \Omega(d_i^c, CP_{cl}^c)^2 \right) \right)} \tag{5.19}$$

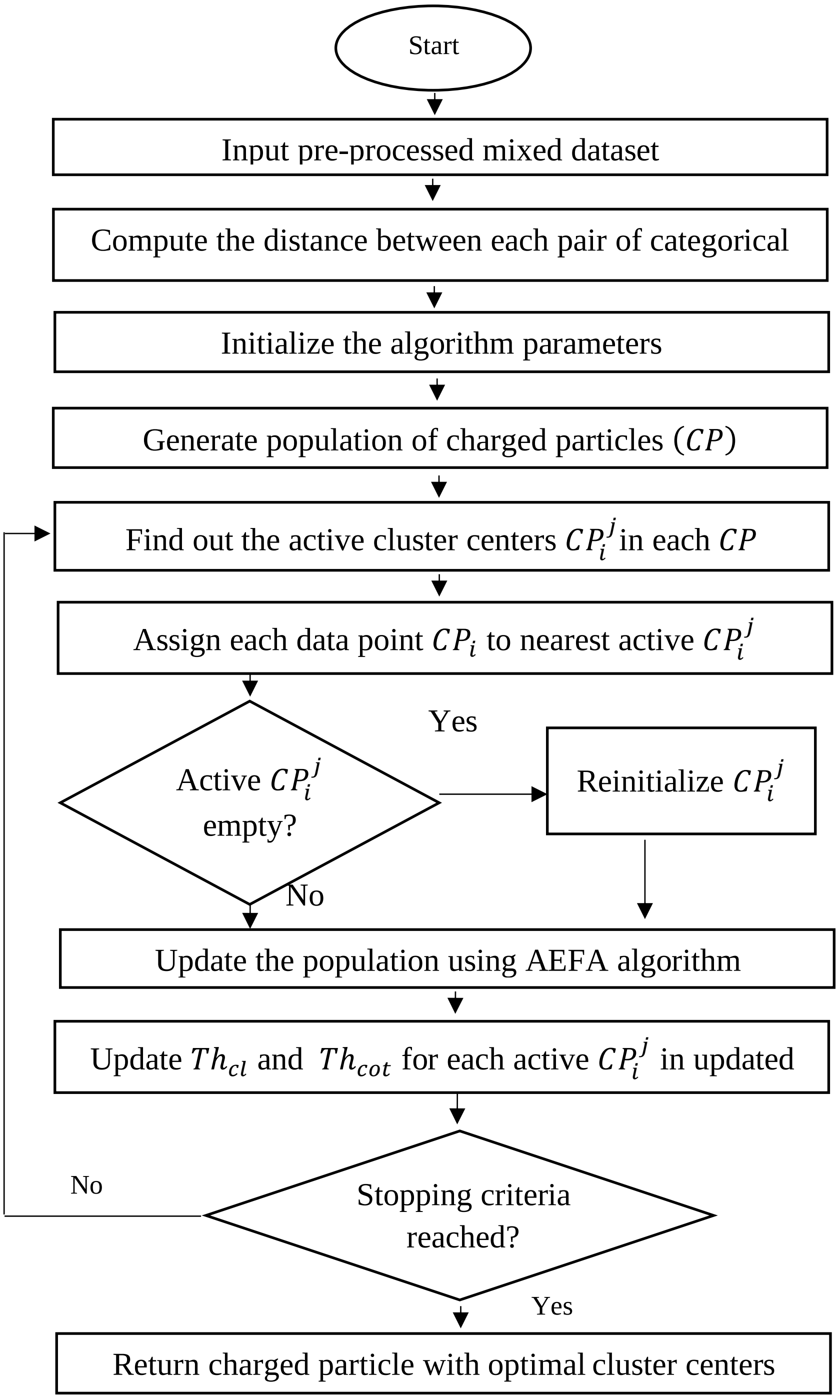


Figure 5.1: Workflow of the proposed clustering algorithm

Where, Th_{cl} represents the selection threshold of a cluster center. $\Omega(d_i^r,\ CP_{cl}^r)$ and $\Omega(d_i^c, CP_{cl}^c)$ represents distance between i^{th} data points and corresponding cluster center of the numeric and categorical attribute, respectively. K_{active} represents number of active cluster centers. n_{cl}represents the total number of data-points belongs to an l^{th}active cluster center.

Algorithm 6: Proposed Clustering Algorithm for the Mixed Dataset

```
Input: Dataset with both mixed numeric and categorical attributes and $n$ data points
Output: Optimal number of cluster centers
1. Define the maximum number of iterations ($N$), the maximum number of clusters ($K_{max}$), population size ($P$), selection threshold ($Th_{cl}$) and cut-off threshold ($T_{cot}$)
2. Compute the dimension ($D_x$) and generate an initial population of charged particles ($CP$) as cluster centers
3. Initialize $iter$ =1
4. while $iter < N$ do
  4.1 for each $i$ in $p$ do
          for each $j$ in $K_{max}$ do
              if ($Th_{ij}$>$T_{cot}$) then
                  Select and activate the cluster center $CP_i^j$ using Eq. (5.17)
              else
                  $CP_i^j$ is set to inactive
              end if
          end for
          for each data point $CP_i$ in given mixed dataset
          (a) Compute the distance between $CP_i$ and active $CP_i^j$ using Eq. (5.4) and assign the $CP_i$ to the closest active $CP_i^j$
          (b) Verify and reinitialize the empty $CP_i^j$ as described in section 5.2.3
          end for
      end for
  4.2 Update the population using the AEFA algorithm (mentioned in section 5.1.3). The exploration process is guided by the fitness function using Eq. (5.20) and Eq. (5.21), and
```

```
            the distance measure for the mixed dataset using Eq. (5.4)
        4.3 Update $Th_{cl}$ and $T_{cot}$ for each $CP$ in the updated population using Eq. (5.18) and Eq.
            (5.19)
        4.4 $iter = iter + 1$
    end while
5. Return the charged particle ($CP$) with optimal cluster centers.
```

5.2.3 Empty Clusters Validation

A cluster center is said to be empty if no data points or less than 2 data points are assigned to it. Such problems are resolved by reinitializing the cluster center for that candidate solution. The candidate solution is reinitialized by assigning n/k_{active} data points to each nearest active cluster center.

5.2.4 Fitness Evaluation

In this step, the performance of the candidate solutions (cluster centers) is measured. As the performance critically relies upon a suitable cluster validation criterion. A random selection of criteria for clustering may lead to poor results. Therefore, the SD index[150] is chosen for cluster validation, and the resulting fitness of the candidate solution is computed as follows:

$$Fitness\ (CP_i) = SD\ (CP_i) * \frac{K_{max} - k_{active}}{k_{active} + 1} \tag{5.20}$$

Where,

$$SD\ (CP_i) = a * Scat(CP_i) + Dist\ (CP_i) \tag{5.21}$$

$Scat(CP)$ and $Dist\ (CP)$ represents intra-cluster distance (compactness) and inter-cluster distance (separation), respectively. $Scat(CP)$ and $Dist\ (CP)$ are computed as follows:

Where, $Scat(CP_i) = \frac{1}{k_{active}} \sum_{i=1}^{k_{active}} \frac{\left\| \sigma_{(CP_i)} \right\|}{\left\| \sigma_{(x)} \right\|}$, $\quad Dist\ (CP_i) \frac{D_{max}}{D_{min}} \sum_{k=1}^{k_{active}} \frac{1}{\left(\sum_{z=1}^{k_{active}} \left\| \vartheta\left(CP_l^k,\ CP_l^z\right) \right\| \right)}$

$$\sigma_{(CP_i)}^{p} = \frac{\sum_{k=1}^{n} \vartheta\left(x_k^D, CP_i^D\right)^2}{n_i} \text{ and } \sigma_{(x)}^{p} = \frac{\sum_{k=1}^{n_i} \vartheta\left(x_k^D,\ x^D\right)^2}{n}$$

$$\text{and } D_{max} = \max_{j \in 1,2,3..C}\left(\|\vartheta(v_i, v_z)\|\right)$$

$$\text{and } D_{min} = min_{j \in 1,2,3..c}\left(\|\vartheta(v_i, v_z)\|\right)$$

Where, $\sigma_{(CP_i)}^{D}$ and $\sigma_{(x)}^{D}$ represents variance of i^{th} cluster and variance of dataset along D^{th} dimension, respectively. CP_i^D and x^D represents cluster center of i^{th} cluster and dataset along D^{th} dimension, respectively. A candidate solution with minimum fitness value is considered a highly accurate solution and is thereby selected as an optimal result.

5.3 Experimental Results of Proposed Clustering Algorithm

This section is further divided into three subsections. Subsection 5.1 gives a performance comparison of the proposed algorithm with the existing mixed data clustering algorithms. Subsection 5.2 discusses the application of the proposed algorithm to the clustering of postoperative surgical records, and Subsection 5.3 presents the statistical significance test of the proposed clustering algorithm.

5.3.1 Performance Comparison of the Proposed and existing Mixed Data Clustering Algorithms

At first, 5 real-life datasets are used to evaluate the performance of the proposed clustering algorithm. These datasets are obtained from the UCI machine learning repository https://archive.ics.uci.edu/ml/datasets.php). The description of the datasets is shown in Table 5.2. Then, the performance of the proposed clustering algorithm is compared with existing mixed dataset clustering algorithms. For performance comparison, two cluster quality measures: average accuracy[121] and standard deviation[121], are used. Average accuracy represents the quality, and the standard deviation represents the reliability of the clustering algorithm. During each iteration of the proposed clustering algorithm, both average accuracy and standard deviation contribute to the robustness of the clustering algorithm, where a high value of average accuracy and lower standard deviation makes a clustering algorithm more robust.

Table 5.2: Characteristics of the used datasets

Dataset	**Data points**	**Attributes** **Numeric**	**Categorical**	**Classes**
Heart Disease (1)	303	5	8	2
Heart Disease (2)	270	6	8	5
Credit Approval	690	6	8	2
Iris	150	4	-	3
Soybean	47	-	35	4

A. Parameter Setting

For experiments, the parameters used in the proposed clustering algorithm, i.e., population size (P), the maximum number of clusters (K_{max}), Coulomb's constant (K_0), and the maximum number of iterations (N) are initialized as follows: P is set to 20, K_{max} is set to $\sqrt{n}$, where n is the size of the dataset, K_0 is set to 500, and N is set to 20.

B. Performance Comparison

The performance of the proposed clustering algorithm is compared with existing mixed data clustering algorithms: K-means clustering algorithm for mixed dataset (KMCMD)[120], and K-harmonic means clustering algorithm for the mixed dataset (KHMCMD)[121], K-prototype (KP) clustering for mixed data[151], improved K-prototype (IKP) clustering algorithm for mixed data[152], SBAC[153], and Ji et al.[154] The experiments are performed in 50 iterations. In every 10 iterations, the number of correct predictions of data points and corresponding class labels obtained by the proposed algorithm is computed in terms of average accuracy (AC). Finally, among the obtained 5 results, the results with the maximum average accuracy (AC) and minimum standard deviation (Sd) of most frequently selected active clusters is chosen as optimum cluster centers. The results are presented in Table 5.4. Table 5.4 demonstrates the comparative results of the AC and Sd among the proposed clustering algorithm and compared algorithms for the used real-life datasets. According to Table 5.4, for Heart Disease (1) dataset with 303 data points containing 5 numeric and 8 categorical attributes, the proposed clustering algorithm has achieved the best AC followed by KHMCMD, KMCMD, IKP, Ji et al., SBAC, and KP. The proposed clustering algorithm shows 6%, 0.7%, 2%, 3.8%, 9.4%, and 26.9 % more accurate results than KHMCMD, KMCMD, IKP, Ji et al., SBAC, and KP, respectively. For Heart Disease (2) dataset with 270 datapoints and having 6 numeric and 8 categorical attributes, the proposed clustering algorithm have produced 1.2%, 2.06%, 17.5%, 28.2%, and 28.3% more accurate results than KHMCMD, KMCMD, IKP, KP, and SBAC, respectively. For the credit approval dataset with 690 datapoints and having 6 numeric and 8 categorical attributes, the proposed clustering algorithm achieved 1%, 3.9%, 8.3%, 30%, and 30.7% more accurate results than KHMCMD, KMCMD, IKP, KP, and SBAC, respectively. For the Iris dataset with 150 data points and having 4 numeric attributes, the proposed clustering algorithm have achieved 9.8%, 10.1%, and 54.7% more accurate results than IKP, KP, and SBAC, respectively. For the soybean dataset with 47 data points and having 35 numeric attributes, the proposed clustering achieved 1%, 5.4 %, and 29.3% more accurate results than IKP, KP, and SBAC, respectively. The results in Table 5.4 demonstrate that for all 5 datasets, the AC of the proposed clustering algorithm is higher. In contrast, the Sd for the proposed clustering algorithm achieves a lower value as compared to the existing mixed data clustering algorithm. The comparative results of Table 4 are presented as a graph in Figure 5.2, which shows the AC of the proposed clustering algorithm and existing clustering algorithms for the used real-life datasets. The height of the bar represents the measured AC, so that the longer the bar, the higher the AC. The Sd obtained for the corresponding dataset is mentioned at the top of the bar. Figure 2 represents that the proposed clustering algorithm achieved the highest AC as well as the lowest Sd for all the datasets. Thus, Figure 5.2 reveals that the proposed clustering algorithm is more robust to optimal cluster center selection.

5.3.2 Application of the Proposed Clustering Algorithm to Postoperative Surgical Patients

Surgical patient clustering is defined as a process of arranging the patients into distinct groups based on their similarity of characteristics. These characteristics include age, gender, body mass index (BMI), American Society of Anesthesiologists (ASA) fitness grade, etc. For multi-specialty hospitals, where an enormous number of patients receive their surgical care, it is quite challenging to manage the surgical records efficiently. Efficiently managed surgical records help hospitals to improve patients care and monitoring to enhance the efficiency of resources within the hospital. In this section, the surgical record management procedure (SRMP) of a multi-specialty hospital, India, is examined. The proposed algorithm is implemented on the hospital's existing SRMP to cluster surgical records and enhance existing SRMP.

A. Dataset Description

The historical postoperative surgical mixed dataset is obtained from Shri Mahant Indiresh Hospital, Dehradun, India. The description of the datasets is given in Table 5.3.

B. Significance of Active Clusters Selected by the Proposed Clustering Algorithm

During each iteration of the algorithm, the significance of the obtained clusters is computed using two parameters: (1) frequency of the number of active clusters selected and (2) average accuracy[121]. The frequency of clusters n_c is computed as:

$$F_Q(n_{c_i}) = \frac{S_{N_{c_i}}}{R_t} \tag{5.22}$$

Where, $S_{N_{c_i}}$ is the number of times a particular cluster is selected, and R_t represents the total number of iterations. Since the surgical dataset has no class label, the average accuracy of the proposed clustering algorithm is measured using SD index. The average accuracy is computed as the inverse of the SD index, and the result with the minimum SD index is considered as a highly accurate clustering outcome.

Table 5.3: Postoperative surgical dataset characteristics

Attribute	**Type**	**Description**
Age	Numeric	Patient's age at the time of surgery
Gender	Categorical	Gender of patient
BMI	Numeric	Body mass index of patient
ASA fitness grade	Numeric	Patients' physical status required before surgery
Marital Status	Categorical	Marital status of patients
Ethnicity	Categorical	Ethnicity of patient
Comorbidity	Numeric	Charlson comorbidity index
Type of Surgery	Categorical	Classifies attempt of surgical procedure
Surgery Duration	Numeric	Length of surgical procedure
Procedural code	Categorical	Primary procedure code
Diagnose code	Categorical	Primary diagnosis code
Surgery Domain	Categorical	Classifies surgical procedure
Grade of Surgery	Categorical	Classifies risk of surgical procedure to the life of the patient
Urgency of surgery	Categorical	Classify the schedule of a surgical patient.
LOS	Numeric	Length of stay in hospital after surgery (in days)

C. Results

The experiments are carried out in 50 iterations, and the results obtained in the pair of 10-10 iterations are shown in Table 5.5. Table 5.5 demonstrates the frequency of selecting the active clusters and the average accuracy (AC) of the clustering outcome determined by the proposed clustering algorithm.

In each pair of iterations, the active clusters' selection frequency and their corresponding SD index value are computed. The average accuracy (AC) and standard deviation (Sd) of the obtained clustering outcomes are computed using the obtained SD index value. Finally, the active cluster with the maximum AC and the minimum Sd is selected as optimum cluster centers. According to Table 5.5, the proposed algorithm selects 6 optimal active clusters most frequently with AC and Sd of 0.77, 0.18, respectively, in the first 10 iterations. In the next 11-20 iterations, 4 optimal active clusters are frequently selected with AC and Sd of 0.81 and 0.23, respectively. During iterations 21-30, 6 optimal active clusters are selected frequently with AC and Sd of 0.86 and 0.14, respectively. In 31- 40 iterations, the proposed clustering algorithm selects 6 optimal active clusters frequently with AC and Sd of 0.87, 0.12, respectively. In the final 41-50 iterations, 5 optimal active clusters are selected frequently with AC and Sd of 0.82 and 0.21. The clustered visualization of iteration-based results, demonstrated in Table 5.5, is presented in Figure 5.3. Figure 5.3 shows the scatter plot of the most frequently selected active clusters during each pair of 10 iterations. Since the dataset has both mixed numeric and categorical attributes, the scatter plots are generated based on the distance between the data points and active cluster centers. Each cluster in the figure is represented by a cluster name, and a circle is used to distinguish between the obtained active cluster partitions. It is too noteworthy to mention here that the circle does not represent the shape of any cluster. Figure 5.3 clearly shows that the proposed clustering algorithm has produced nonoverlapping and well-separated cluster partitions in every pair of iterations. Further, due to higher average accuracy and lower standard deviation, the clusters in Figure 5.3(d) are considered optimal cluster partitions. The overall performance is summarized in Table 5.6 and plotted as line graph in Figure 5.4. It shows that the proposed clustering algorithm obtains 6 optimal active clusters with a frequency of 1.2 and an AC of 0.87 with Sd of 0.12.

Table 5.4: The performance comparison of proposed and existing mixed data clustering algorithms

Dataset	**Proposed Algorithm**		**Improved K-Prototype**		**SBAC**		**KHMCMD**		**KMCMD**		**Ji et al.[154]**	
	AC	***Sd***	***AC***	***Sd***	***AC***	***Sd***	***AC***	***Sd***	***AC***	***Sd***	***AC***	***Sd***
Heart Disease (1)	0.846	0.14	0.826	~	0.577	~	0.840	0.15	0.8389	0.15	0.808	~
Heart Disease (2)	0.828	0.18	0.653	~	0.546	~	0.816	0.33	0.8074	1.20	~	~
Credit approval	0.862	0.11	0.779	~	0.562	~	0.852	0.38	0.8223	12.77	0.794	~
Iris	0.92	0.18	0.822	~	0.819	~	~	~	~	~	~	~
Soybean	0.91	0.16	0.90	~	0.856	~	~	~	~	~	~	~

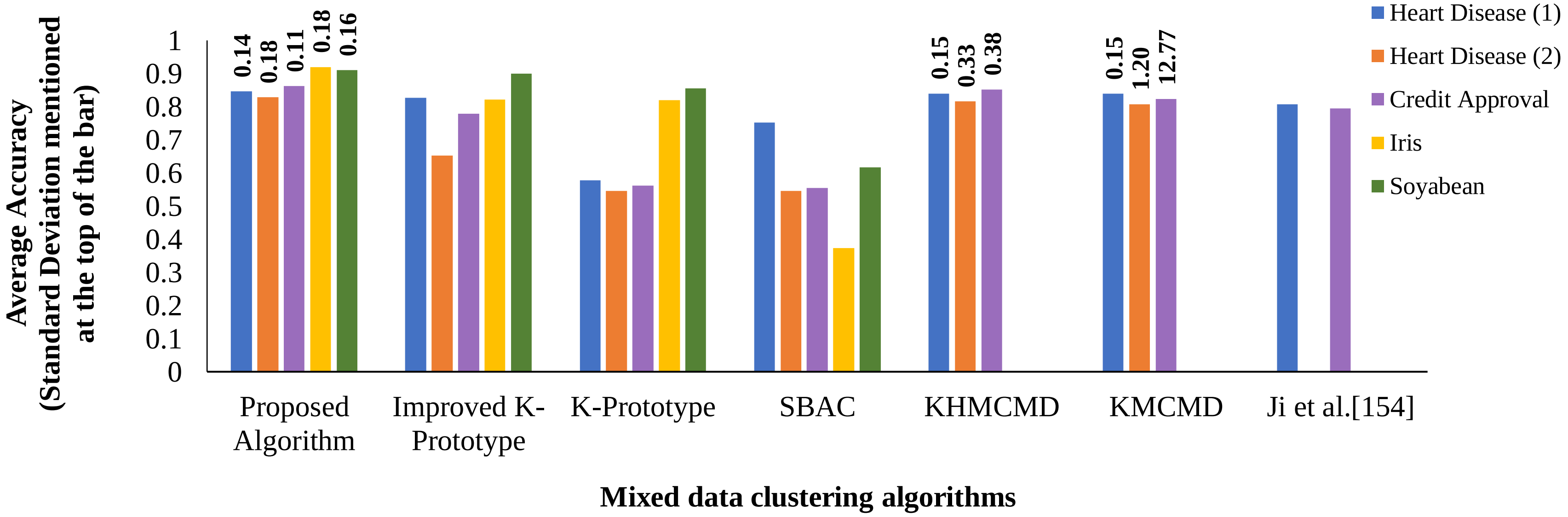
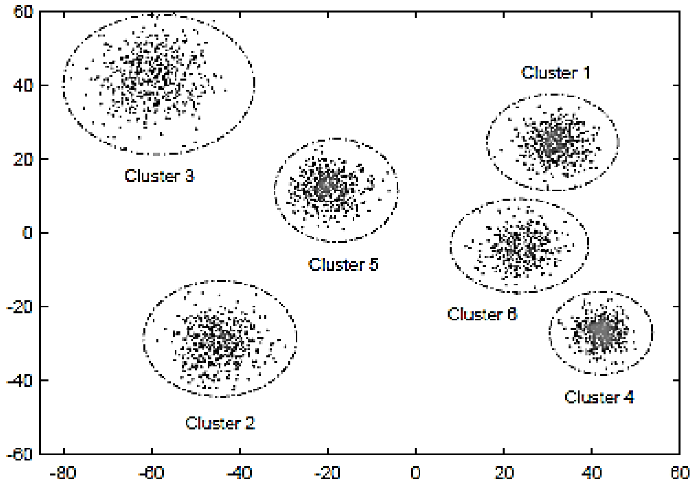


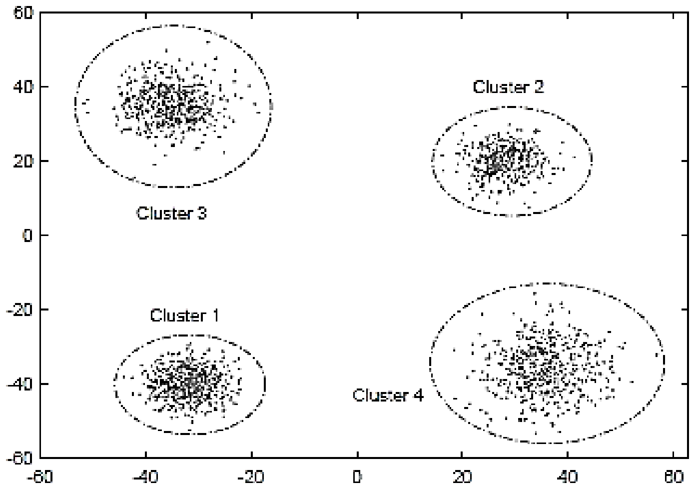
Figure 5.2: Comparison of clustering accuracy of proposed and existing clustering algorithms

Table 5.5: Frequency (F_Q) and average accuracy (AC) (and standard deviation (Sd)) of the clusters selected by the proposed clustering algorithm for the postoperative surgical dataset

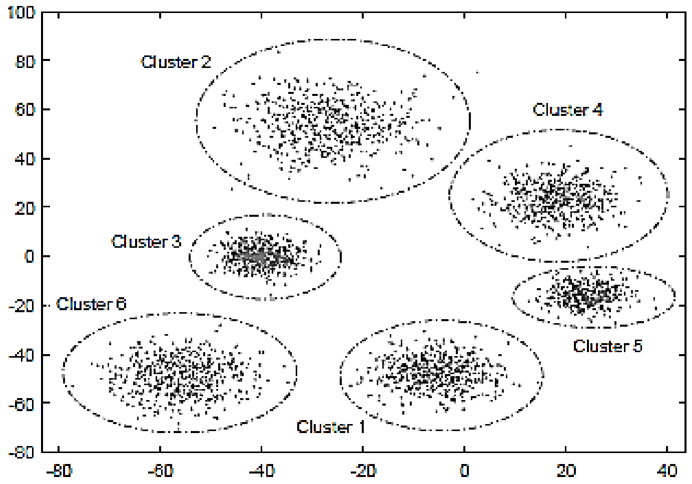
Iterations		**No of active cluster extraction**						
		2	**3**	**4**	**5**	**6**	**7**	**8**
1-10	**F_Q**	0.0	0.5	0.8	0.4	**1.4**	0.6	0.3
	AC	0.0	0.51 (±0.63)	0.62 (±0.42)	0.47 (±0.360)	**0.77 (±0.18)**	0.59 (±0.48)	0.44 (±0.44)
11-20	**F_Q**	0.4	0.5	**0.9**	0.6	0.6	0.8	0.2
	AC	0.46 (±0.39)	0.54 (±0.34)	**.81 (±0.23)**	0.58 (±0.42)	0.69 (±0.24)	0.75 (±0.21)	0.33 (±0.56)
21-30	**F_Q**	0.0	0.0	0.8	0.9	**1.5**	0.2	0.6
	AC	0.0	0.0	0.68 (±0.35)	0.78 (±0.21)	**.86 (±0.14)**	0.37 (±0.64)	0.51 (±0.60)
31-40	**F_Q**	0.3	0.0	0.9	1.0	**1.2**	0	0.6
	AC	0.32 (±0.69)	0.0	0.54 (±0.43)	0.62 (±0.31)	**0.87 (±0.12)**	0.0	0.41 (±0.48)
41-50	**F_Q**	0	0.8	0.8	**1.1**	0.6	0.7	--
	AC	0.0	0.46(±0.36)	0.53 (±0.62)	**.82 (±0.21)**	0.39 (±0.46)	0.48 (±0.54)	--



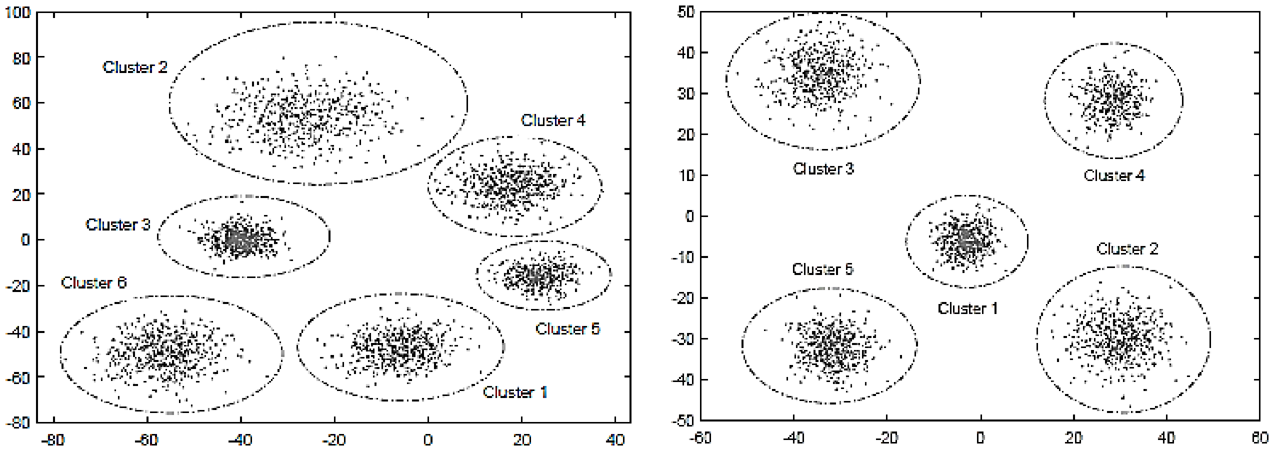
(a) Iterations (1-10): Clusters = 6, F_Q= 1.4, AC= 0.77, Sd=0.18



(b) Iterations (11-20): Clusters = 4, F_Q= 0.9, AC =0.81, Sd= 0.23



(c) Iterations (21-30): Clusters = 6, F_Q= 1.5, AC = 0.86, Sd= 0.14



(d) Iteration (31-40): Cluster = 6, F_Q= 1.2, AC= 0.87, Sd= 0.12

(e) Iteration (41-50): Cluster = 5, F_Q= 1.1, AC= 0.82, Sd= 0.21

Figure 5.3: Cluster center selection frequency (F_Q), average accuracy (AC) and standard deviation (Sd) for the postoperative surgical dataset

Table 5.6: Performance of the proposed mixed dataset clustering algorithm on the postoperative surgical dataset

Parameter	**Value**
No of active cluster obtained (standard deviation)	6.0
Frequency of active cluster selection (F_Q)	1.2
Average accuracy (AC)	0.87
Standard deviation (Sd)	$\pm$ 0.12

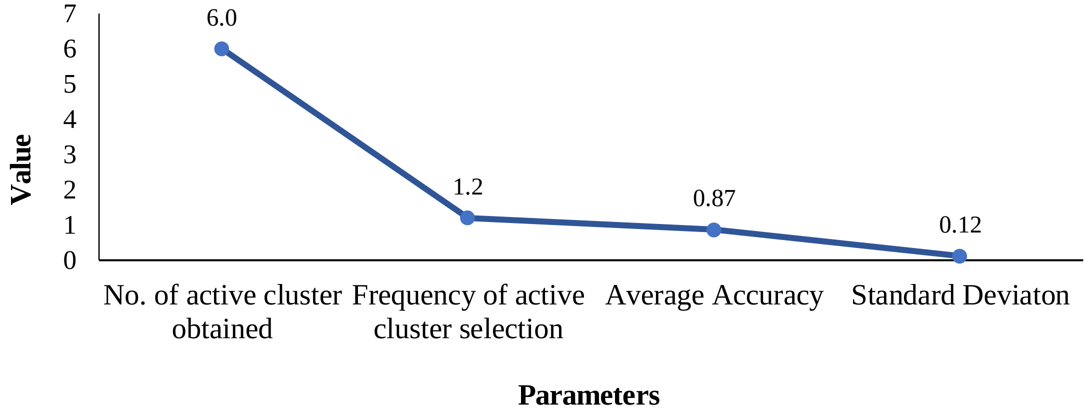


Figure 5.4: Performance of the proposed clustering algorithm on the postoperative surgical dataset

5.3.3 Statistical Significance Test of Proposed Clustering Algorithm

Although results based on average accuracy (AC) and standard deviation (Sd) from Table 5.3 indicate that the proposed clustering algorithm outperformed compared algorithms, the statistical analysis is performed to validate the proposed clustering algorithm's performance. The statistical analysis helps in determining whether the results obtained by the proposed clustering algorithm are significant (P-value < 0.05) or not. In recent years, several methods[148] are used to compare and validate the performance of clustering algorithms. Since the proposed clustering algorithm aims to find optimal cluster centroids (means), a statistical analysis method based on "mean"' knows as unpaired t-test[41], is used to validate the proposed clustering algorithm. The unpaired t-test compares the means of results produced by the proposed clustering algorithm and the second-best performing clustering algorithm (KHMCMD) of Table 5.4. The results obtained in Table 5.7 demonstrate that P-value satisfies the significance level (< 0.05) for all datasets, which shows the clustering results produced by the proposed clustering algorithm are statistically significant than the other existing clustering algorithm.

Table 5.7: Unpaired t-test between the proposed and existing KHMCMD clustering algorithm for real-life datasets

Dataset	**Standard Error**	***t***	**95% Confidence Interval**	**Two-tailed P-value**	**Significance**
Heart Disease (1)	1.000	10.0000	-12.306 to -7.693	0.0001	Extremely Significant
Heart Disease (2)	1.811	13.0309	-27.78 to -19.42	0.0001	Extremely Significant
Credit Approval	1.00	5	-7.306 to -2.693	0.0011	Significant
Iris	1.87	4.06	-11.914 to -3.285	0.003622	Significant
Soybean	1.522	3.8335	-9.28 to -2.39	0.0040	Significant
Postoperative surgical dataset	1.709	2.9450	-9.0725 to -0.9916	$3.2e^{-5}$	Significant

Chapter 6

OPTIMAL SURGICAL TEAM SELECTION

Addressing the surgical team selection challenge faced by multispecialty hospitals, a decision-making model is proposed in this chapter to select an optimal list of surgical teams for a patient. The proposed model deals with two critical issues: To improve the existing surgical history management system by arranging surgical patients into optimal sub-groups based on their similar characteristics, and 2) based on surgical patient sub-groups, to select an optimal list of surgical teams for a new surgical patient. For arranging surgical patients into optimal sub-groups and optimal surgical team selection, a population-based meta-heuristic clustering algorithm for the mixed dataset and a population-based meta-heuristic algorithm for multi-objective optimization are proposed in this chapter.

6.1 Multi-Objective Optimization

A multiple-objective problem (MOOP) can be a minimization problem or a maximization problem. It involves O distinct target objectives defined as follows:

Minimize / Maximize:

$$\text{Fitness}\ (P_r) = \ [\text{Fitness}_i(P_r),\ i = 1, 2 \dots O] \tag{6.1}$$

Subject to constraints: $EC_j(P_r) \leq 0,\ j = 1,2,..j$, and $IC_k(P_r)\ \leq 0,\ k = 1,2, \dots k$

Where, $\text{Fitness}_i(P_r)$ represents i^{th} objective function of P_r^{th} solution, and $EC_j(P_r)$ and $IC_k(P_r)$ represents j^{th} equality and k^{th} inequality constraints, respectively. In MOOP, the Pareto dominance theory[143] is utilized to determine optimal solutions in global search space, is defined as follows:

According to MOOP, a solution vector $\vec{l}$ =$l_1;\ l_2;\ \dots;\ l_d$ shows dominated by vector $\vec{r}$ = $r_1;\ r_2;\ \dots r_d$, if and only if

$$\forall\, i \in \{1 \dots. d\}, r_i \le l_i \;\exists\, i \in \{1 \dots. d\} : r_i < l_i$$

Where, d implies the dimension of the objective space. If a solution $\vec{r} \in R$ is not dominated by any other solution ($l \in L$), it is called Pareto optimal. These solutions ($\vec{r}$) are said to be a non-dominated set of solutions, also called Pareto optimal set.

6.2 Proposed Surgical Team Selection Model

This section provides a detailed description of the proposed model. The proposed model is comprised of two modules; surgical history management (SHM) and surgical team selection (STS). The SHM module performs two activities: 1) clustering of existing surgical patients based on their characteristics, and 2) the filtration of existing surgical team details. The STS module produces an optimal list of surgical teams for a given patient. The SHM module is designed to assist the STS module in decision-making. The workflow of the proposed model is presented in Figure 6.1.

6.2.1 Surgical History Management (SHM) Module

Surgical history is a vital aspect of medical records and includes social and demographic records of surgical patients, surgical team details, and diagnostic &

procedural test. For multispecialty hospitals that provide surgical services to numerous patients, efficient management of surgical records is essential. An efficiently organized surgical history help hospitals to enhance their patient care and resource efficiency. To utilize these surgical records, the following two activities are performed in the SHM module.

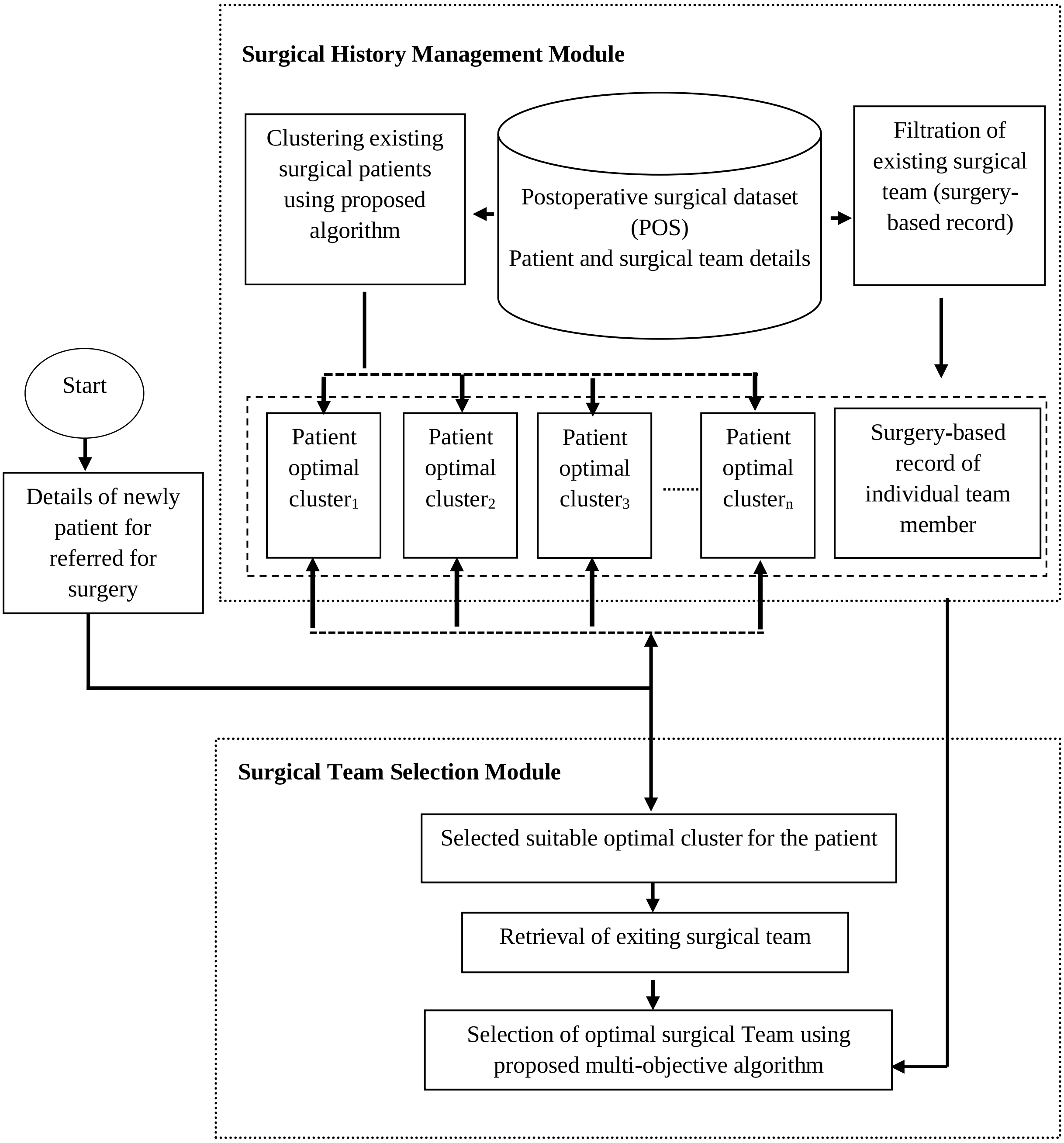


Figure 6.1: Model for selection of optimal surgical team

A. Clustering of Existing Surgical Patients

Arranging existing surgical patients in disjoint clusters based on their characteristics helps hospitals to find a possible sub-group for a newly referred surgical patient. Therefore, in this section, an efficient data clustering algorithm (discussed in chapter 5) for the mixed dataset is proposed to cluster surgical patients. The proposed clustering algorithm (Algorithm 7) based on the AEFA (discussed in chapter 5), focuses on finding optimal clusters automatically. The steps performed in the proposed clustering algorithm are as follows:

Algorithm 7: Proposed Algorithm for Clustering of Surgical Patients

---

```
Input: A postoperative surgical dataset
Output: Patients and their optimal clusters
1. Define the maximum number of iterations ($MaxIT$), the maximum number of
   clusters ($CLC_{max}$), population size ($P_S$), selection threshold ($THV_{SL}$) and cut-off
   threshold ($T_{COV}$)
2. Compute the dimension ($Dim_{QP}$) and randomly initialize a population of patients
   ($PT$) as cluster centroids from the dataset
3. Initialize iteration counter, $I_t$ =1
4. while $I_t < MaxIT$ do
   4.1 for each $i$ in $P_s$ do
           for each $j$ in $CLC_{max}$ do
               if ($THV_{ij}$>$T_{COV}$) then
                   Verify and activate the centroid $PT_i^j$ using Eq. (6.6)
               else
                   $PT_i^j$ is set to inactive
             end if
           end for
           for each patient ($PT_i$) in the given mixed dataset do
```

```
            (a) Compute the distance between $PT_i$ and active $PT_i^j$ using the
                distancemeasures (discussed in chapter 5) and assign the $PT_i$ to the
                nearest active $PT_i^j$
            (b) Verify and reinitialize the empty $PT_i^j$ as described in chapter 5
        end for
    end for
    4.2 Update the population using the AEFA algorithm (sub-section 6.2.1). The
        fitness function in Eq. (6.8) and the distance measures (discussed in chapter 5)
        are used to direct the exploration process.
    4.3 Update $THV_{SL}$ and $T_{COV}$ for each $PT$ in the updated population using Eq. (6.7)
    4.4 $I_t = I_t + 1$
    end while
5. Return the patients $(PT)$ and their optimal clusters.
```

(i) Improved Electrostatic Force Computation and Velocity Update

In traditional AEFA[144], total electrostatic attraction force (F_i^D) (Eq. 6.2) on a i^{th} charged particle is computed by multiplying a random number to the electrostatic force exerted by j^{th} charged particles on it. This force affects the acceleration and velocity, thus resulting in the global movement of all charged particles. Furthermore, the velocity (Eq. 6.4) of i^{th} charged particle is updated by multiplying a random number to its existing velocity value. These random numbers add stochastic behavior in the algorithm's searching process, resulting in an imbalance between exploration and exploitation, thus causes the algorithms to trap in local optima. To maintain a balance between exploration and exploitation, instead of using only random number we have considered charge of a charged particle (q_i) also. The charge of a charged particle controls the stochastic behavior during the computation of F_i^D and velocity; this, in turn, reduces the acceleration and velocity values of the charged particle, thus balancing exploration and exploitation. The modified equations (Eqs. 2 and 4) are as follows:

$$F_i^D(T) = \sum_{j=1, j\neq i}^{N} rand_i\,() * F_{ij}^D(T) \tag{6.2}$$

$$F_i^D(T) = \sum_{j=1, j\neq i}^{N} (rand_i \; * r_1 + (1 - e^{\sqrt{q_i}}) * r_2) * F_{ij}^D(T) \tag{6.3}$$

$$vel_i^D(T+1) = \; rand_i * \; vel_i^D(T) + a_i^D(T) \tag{6.4}$$

$$vel_i^D(T+1) = \left(rand_i * r_1 + \left(1 - e^{\sqrt{q_i}}\right) * r_2\right) * \; vel_i^D(T) + a_i^D(T) \tag{6.5}$$

Where, r_1 and r_2 are two non-negative integers and r_1 + r_2 =1. Further, in case of r_1=1 and r_2= 0, Eq. (6.2) & Eq. (6.4), and Eq (6.3) & Eq. (6.5) are treated identically.

(ii) Selection of Active Centroids

For each candidate solution, active centroids are selected from CLC_{max} centroids based on the following condition:

$$CLC_{ij} = \begin{cases} 1, \; THV_{ij} > T_{COV} \\ 0, \; Otherwise \end{cases} \tag{6.6}$$

Where, T_{COV} is the cut-off value for every centroid, and it is set to a random value between [0,1]. T_{COV} depends on the selection threshold value (THV_{SL}) of a centroid and computed as follows:

$$T_{COV} = \frac{1}{CLC_{active}} \sum_{l=1}^{CLC_{active}} THV_{SL} \tag{6.7}$$

Where, $$THV_{SL} = \sqrt{\frac{1}{n_{SL}}((\sum_{i=1}^{m_r} \Omega(P_i^r,\ CLC_{SL}^r)^2) + \left(\sum_{i=1}^{m_c} \Omega(P_i^c, CLC_{SL}^c)^2\right)}$$

Where, CLC_{active} represents the number of active centroids in each candidate solution.

(iii) Computation of Fitness

The efficiency of a clustering algorithm depends upon the cluster validation criteria. So, the Silhouette index (SI)[148] is used for cluster validation. The fitness of the candidate solution is computed as follows:

$$Fitness(QP_i) = SI(QP_i) * \frac{CLC_{max} - CLC_{active}}{CLC_{active} + 1} \tag{6.8}$$

Where, $$SI(QP_i) = \frac{Mean_g - Mean_h}{max\ (Mean_g, Mean_h)}$$

Where, $Mean_g$ represents the mean distance to other data points in the same cluster, i.e., mean intra-cluster distance, and $Mean_h$ represents the mean distance to the other data points in different clusters, i.e., mean inter-cluster distance. A candidate solution, having minimum fitness, is selected as an optimal result.

B. Filtration of the existing Surgical Team

In this section, a postoperative surgical dataset is considered as an input. Subsequently, on the basis of the required surgery type (e.g., orthopedic surgery, neurosurgery, and pediatric surgery etc.), the details of existing surgical teams are retrieved. A surgical team involves a surgeon, a nurse circulator, and an anesthesiologist. For each retrieved surgical team, additional information such as the complication rate and patient's surgical feedback rating are computed and stored in a database. This stored information helps decision makers in optimizing the process of surgical team selection.

6.2.2 Surgical Team Selection (STS) Module

This module is invoked when the proposed model receives details of a new surgical patient. Subsequently, an optimal cluster is selected for the new patient. The details of the corresponding surgical teams are then retrieved from the selected cluster and processed to obtain the optimal list of surgical teams. In this section, an efficient meta-heuristic algorithm for multi-objective optimization is proposed to produce an optimal list of surgical teams.

A. Proposed Multi-Objective Optimization Algorithm (MOOA) for Surgical Team Selection

The proposed MOOA algorithm begins with parameter initialization. Subsequently, the population of candidate solutions is generated. Finally, based on the fitness value of each candidate solution, the best solution is selected. This process is performed iteratively until the convergence condition is satisfied, and finally, optimum solutions are obtained. Further, to improve convergence and to prevent solution from trapping in local optima, bounded exponential crossover (BEX)[153] and polynomial mutation operator (PMO)[154] are used in the proposed algorithm. The proposed algorithm is presented in Algorithm 8.

(i) Population Initialization

The proposed algorithm has two populations; search population (P_{Search}) and external population ($P_{External}$). The P_{Search}, which contains initial candidate solutions, computes the non-dominant solutions and stores them in the $P_{external}$. The surgical

Algorithm 8: Proposed Multi-Objective Optimization Algorithm for Surgical Team Selection

```
Input: Details of the existing surgical teams (From STS and SHM modules)
Output: Optimal list of surgical Teams
1. Define search population size ($P_{SSize}$), external population size ($P_{ESize}$), and
   maximum no. of iterations ($MaxIT$)
2. Initialize search population ($P_{Search}$) of the surgical team (S$T$) obtained from
   the STS module.
3. Initialize external population $P_{External} = \emptyset$, and set iteration counter $I_t = 1$.
4. while ($I_t$< $MaxIT$)
   4.1 for each $ST \in P_{Search_{I_t}}\ U\ P_{External_{I_t}}$
       (a) Compute the fitness $Fitness(ST)$ using Eqs. (6.10) - (6.11)
       (b) Compute the additional density value (Density) using $k^{th}$ nearest
           neighbor algorithm.
       end for
   4.2 for each ST ∈ $P_{Search_{I_I}}\ U\ P_{External_{I_t}}$ do
          if $Fitness(ST) < 1$ then
             $P_{External_{I_{t+1}}} = P_{External_{I_{t+1}}}\ U\ \{ST\}$
          end if
       end for
   4.3 if (size of$\left(P_{External_{I_{t+1}}}\right) < P_{ESize}$) then
          $P_{External_{I_{t+1}}} = P_{External_{I_{t+1}}}\ U\ \left(\begin{matrix}\left(P_{Search_{I_t}}\ U\ P_{External_{I_t}}\right) \\ \left[1{:}\,P_{ESize}\left|P_{External_{I_{t+1}}}\right|\right]\end{matrix}\right)$
       else
          Compute the additional density values for each non-dominated $ST$ in
          $P_{External_{I_{t+1}}}$delete a S$T$ with the smallest density values
       end if
```

```
    4.4 Select surgical team (ST) into the mating pool $P_{I_{t+1}}$ from
        $P_{Search_{I_t}}\ U\ P_{External_{I_{t+1}}}$
    4.5 Evaluate charge, update the velocity and position of $POP_{I_t+1}$ and obtain the
        new $STs$
    4.6 Apply crossover and mutation operator (BEX) and (PMO) on population $P_{I_{t+1}}$
    4.7 for each $ST \in P_{External_{I_{t+1}}}$ do
            Compute the additional density value of each ST
        end for
    4.8 for each $ST \in P_{External_{I_{t+1}}}$ do
            Compute the additional density value of each ST in $P_{External_{I_{t+1}}}$
            and perform non- dominated sorting
        end for
    4.9 $I_t = I_t + 1$
  end while
5. Return the optimal list of the surgical teams (ST)
```

team retrieved in the STS module serves as an initial P_{Search}, and along with the surgical team extracted in the SHM module, it is used to conduct the exploration process of the proposed algorithm. The maximum size of the initial population is computed as follows:

$$P_{SSize} = \sqrt{P_{Comb}} \tag{6.9}$$

Where, $\sqrt{P_{Comb}}$ is the number of possible combinations of surgical team extracted from the selected suitable patient cluster.

(ii) Fitness Evaluation

In this section, two objectives are considered to evaluate the performance of a surgical team: complications associated with surgery[55] and patient's surgical feedback rating. Surgical feedback is defined as the experience sought by the patient during the surgical period. Feedback ratings of the existing surgical teams are collected in terms of the surgical team's behavior and activity. The fitness functions are computed using as follows:

Minimize: $$Fitness(ST) = CompF_t(ST) \tag{6.10}$$

and Maximize: $$Fitness(ST) = SF_t(ST)$$

Where, $$CompF_t(ST) = \alpha * Comp_t(ST) + (1-\alpha) * noComp_t(ST)$$

Where, $\mathrm{CompF}_t(ST)$[55] is a fitness value that represents a combination of complication ratio $Comp_t$, and no-complication ratio $noComp_t$ associated with the surgical team (ST) at time t. SF_t represents a surgical feedback rating. It is computed as follows:

$$SF_t = \frac{1}{n}\sum_{i=1}^{n}\frac{\sum_{J=1}^{SP}\overline{Rating}_{ST_{ji}}}{SP_{ST_i}} \tag{6.11}$$

Where, $\overline{Rating}_{ST_{ji}}$ represents the mean of feedback rating provided by j^{th} patient of i^{th} team member. SP_{ST_i} and n are the number of surgical patients treated by i^{th} team member and the total number of team members, respectively. Thus, the overall objective is to find surgical teams with a lower complication ratio and higher feedback ratings.

Because the definition of charge in conventional AEFA was not found suitable to solve MOOP[146], thus this section uses multi-objective function given by SPEA2[157] as the fitness function of AEFA.

$$Fitness(ST_i) = Raw\_Fitness(ST_i) + Density(ST_i),$$

Where, $$Raw\_Fitness(ST_i) = \sum_{j \in \mathrm{POP}_{S_t} \mathrm{U}\, \mathrm{POP}_{ext_t}} dom(j)$$

and $$dom(j)\left|\{i \middle| i \in \mathrm{POP}_{S_t} \mathrm{U}\, \mathrm{POP}_{\mathrm{ext}_i}\}\right|$$

Where, $Fitness(ST_i), Density(ST_i),$ and $Raw\_Fitness(ST_i)$ represent the fitness value, additional density value, and raw fitness value of the i^{th} surgical team, respectively. For each surgical team, $Raw\_Fitness(ST_i)$ is computed in terms of $\mathrm{Comp}_t(ST)$ and $SF_t(ST)$, and it exhibits the strength of a solution (surgical team) by computing the number of other solutions to whom it dominates and assigns a rank to the solution. In a situation where multiple solutions are non-dominant and are assigned similar ranks, the additional density value ($Density$) is used to differentiate various solutions. In this section, k^{th} nearest neighbor algorithm is used to estimate the density.

6.3 Experimental Results of Proposed Surgical Team Selection Model

Experiments were performed using a real case study of a multispecialty hospital in India. The proposed model was implemented using a postoperative surgical dataset (POS) (discussed in chapter 5), which was obtained from the orthopaedic surgery department of the hospital. Sub-section 5.1 discusses the performance of the SHM module. Subsection 5.2 discusses the performance of the STS module. Table 6.1 lists symbols used in the proposed framework.

Table 6.1: Symbols used in the proposed surgical team selection model

Symbol	**Definition**
P_{SSize}	Size of the initial search population
P_{Search}	Search population
P_{ESize}	Size of the external population size
$P_{External}$	External population
QP_i	Position of the i^{th} charged particle (candidate solution)
CLC_{max}	Maximum number of cluster centroids
THV_{SL}	Selection threshold value
CLC_{active}	Number of active cluster centroids
ϑ	Distance between data points and cluster centroids
P_{it}^{r}	t^{th} numeric attribute value of P_i
P_{it}^{c}	t^{th} categorical attribute value of P_i
D	Objective space dimension
P_{Comb}	Number of possible combinations of the surgical team
vel_i^D	Velocity of a i^{th} charged particle in the D^{th} dimension
SF_t	Surgical feedback rating
$\overline{Rating}_{ST_{ji}}$	Mean of surgical feedback rating provided by the j^{th} patient of the i^{th} team member
SP_{ST_i}	The total number of surgical patients treated by the i^{th} surgical team
$CompF_t$	Combination of the complication ratio ($Comp_t$) and the no-complication ratio ($noComp_t$)
$Fitness\ (ST_i)$	Fitness of the i^{th} surgical team
$Density\ (ST_i)$	Additional density value of the i^{th} surgical team
$Raw_{Fitness}(ST_i)$	The raw fitness value of the i^{th} surgical team
$Mean_g$	It represents the average distance to other data points in the same cluster
$Mean_h$	It represents the average distance to the data points of different clusters.
T_{COV}	Cutoff threshold value
$MaxIT$	Maximum number of iterations

$TEAF_i^D$	Total electrostatic attraction force on a i^{th} charged particle in the D^{th} dimension.
$a_i^D(T)$	Acceleration of a i^{th} charged particle in the D^{th}dimension
q_i	Charge on a i^{th} charged particle

6.3.1 Performance Evaluation of Surgical History Management (SHM) Module

The performance of the SHM module was evaluated in three steps. Firstly, the performance of the proposed clustering algorithm was measured using nine real-life datasets (Table 6.2). Secondly, it was compared with six existing clustering for non-mixed dataset: (i) PSO, (ii) hybrid atom search optimization (ASO) and PSO (ASOPSO), (iii) ASO, (iv) hybrid PSO and gravitational search algorithm (PSOGSA), (v) hybrid PSO and firefly algorithm (PSOFA), (vi) hybrid ASO and sine-cosine algorithm (ASOSCA) [158]. The results revealed that the proposed clustering algorithm outperformed existing algorithms (Table 6.3). Subsequently, the performance of the proposed clustering algorithm was also compared with five existing clustering algorithms for mixed dataset: (i) k-means clustering for mixed dataset algorithm[120], (ii) KHMCMD[121], (iii) k-prototypes clustering algorithm[151], (iv) Improved k-prototypes clustering algorithm[152], (v) algorithm proposed by Ji et al.[154]. The comparative results are shown in Table 6.4. Thirdly, the performance of the proposed clustering algorithm was evaluated using the POS dataset. The results revealed that considering all iterations, six active patient clusters with a selection frequency of 1.6, an average fitness of 0.96, and a standard deviation of 0.13 were selected as an optimal solution (Table 6.5).

Table 6.2: Characteristics of real-life datasets

Dataset	**Data points**	**Attributes**			**Classes**
		Numeric	**Categorical**	**Others**	
Breast Tissue	106	9	-	-	6
CMC	1473	2	4	3	2
Wine	178	13	-	-	3
Iris	150	4	-	3	3
Ecoli	336	7	-	-	8
Heart Disease (1)	303	5	8	-	2
Heart Disease (2)	270	6	8	-	5
Credit Approval	690	6	8	-	2
Soybean	47	-	35	-	4

Table 6.3: Comparison of the performance of proposed and existing clustering algorithms using non-mixed datasets

Dataset	**Index**	**Algorithms**						
		PSO	**ASOPSO**	**ASO**	**PSOGSA**	**PSOFA**	**ASOSCA**	**Proposed Algorithm**
Breast Tissue	Silhouette Index	0.74	0.71	0.29	0.77	0.69	0.77	0.82
	Dunn Index	0.43	0.26	0.17	0.52	0.31	0.66	0.71
	Davies-Bouldin index	0.61	0.51	1.07	0.57	0.56	0.63	0.44
CMC	Silhouette Index	0.25	0.17	0.20	0.21	0.22	0.247	0.38
	Dunn Index	0.08	0.07	0.06	0.10	0.09	0.04	0.12
	Davies-Bouldin index	0.61	0.78	0.61	0.61	0.68	0.31	0.30
Wine	Silhouette Index	0.29	0.32	0.23	0.36	0.37	0.52	0.55
	Dunn Index	0.07	0.06	0.05	0.08	0.07	0.12	0.14
	Davies-Bouldin index	0.52	0.56	0.60	0.41	0.43	0.12	0.12
Ecoli	Silhouette Index	0.001	0.06	0.13	0.05	0.00	0.19	0.20
	Dunn Index	0.05	0.06	0.07	0.06	0.06	0.11	0.12
	Davies-Bouldin index	0.69	0.80	0.82	0.72	0.7	0.76	0.51

Table 6.4: Comparison of the performance of proposed and existing clustering algorithms using mixed datasets

Dataset	Proposed Algorithm	Improved K-prototypes	K-prototypes	KHMCMD	KMCMD	Ji et al. [154]
	AC (STD)	AC (STD)	AC (STD)	AC (STD)	AC (STD)	AC (STD)
Heart Disease (1)	0.853 (±0.13)	0.826 (~)	0.577 (~)	0.840 (±0.15)	0.838 (±0.15)	0.853 (±0.13)
Heart Disease (2)	0.830 (±0.19)	0.653 (~)	0.546 (~)	0.816 (±0.33)	0.807 (±1.20)	0.830 (±0.19)
Credit Approval	0.864 (±0.11)	0.779 (~)	0.562 (~)	0.852 (±0.38)	0.822 (±12.77)	0.864 (±0.11)
Iris	0.95 (±0.17)	0.822 (~)	0.819 (~)	~	~	0.95 (±0.17)
Soybean	0.93 (±0.17)	0.90 (~)	0.856 (~)	~	~	0.93 (±0.17)

AC: Average accuracy, STD: Standard deviation. (~ shows results not available)

Table 6.5: Performance evaluation of the proposed surgical patients clustering algorithm on the POS dataset

Number of Active clusters Selected	**Parameters**	**Iterations**				
		1-10	**11-20**	**21-30**	**31-40**	**41-50**
2	Selection Frequency	0.0	0.3	0.0	0.5	0.0
	Average Fitness	0.0	0.70(±0.42)	0.0	0.52(±0.33)	0.0
3	Selection Frequency	0.4	0.3	0.4	0.0	0.6
	Average Fitness	0.74(±0.56)	0.71(±0.52)	0.58(±0.61)	0.0	0.66(±0.25)
4	Selection Frequency	0.6	0.8	0.4	0.6	0.6
	Average Fitness	0.65(±0.40)	0.83 (±0.34)	0.76(±0.39)	0.80(±0.36)	0.52(±0.64)
5	Selection Frequency	0.3	0.7	0.6	0.8	1.0
	Average Fitness	0.36(±0.47)	0.60(±0.36)	0.69(±0.23)	0.55(±0.37)	0.91(±0.22)
6	Selection Frequency	1.0	0.8	1.2	**1.6**	0.4
	Average Fitness	0.62 (±0.16)	0.74(±0.35)	0.92 (±0.18)	**0.96 (±0. 13)**	0.50(±0.32)
7	Selection Frequency	0.4	0.7	0.5	0.5	0.4
	Average Fitness	0.57(±0.58)	0.67(±0.40)	0.30(±0.28)	0.40 (±0.40)	0.52(±0.59)
8	Selection Frequency	0.3	0.1	0.3	0.0	0.0
	Average Fitness	0.53(±0.32)	0.22(±0.38)	0.50(±0.41)	0.0	0.0

6.3.2 Performance Evaluation of Surgical Team Selection (STS) Module

The performance of the STS module of the proposed MOOA is evaluated on the basis of parameters listed in Table 6.6 using three benchmark functions, namely SCH, FON, and ZDT1[147]. Subsequently, the performance of the proposed MOOA was compared with four existing MOOAs: SPGSA[143], NSGA II[147], NSPSO[159], and BCMOA[160], on the basis of three performance parameters, namely converge metric [CM], diversity metric [DM][147], and generational distance metric (GD)[149]. A minimum value of all these parameters is desired for optimal solutions. The results demonstrated that the proposed algorithm achieved a minimum value for CM, DM, and GD (in terms of the mean) for all considered benchmark functions (Table 6.7). This finding indicated that the proposed algorithm outperformed existing MOOAs in terms of the convergence rate while maintaining the diversity among optimal solutions. The results shown in Table 6.7 are presented as a graph in Figure 6.2. For better representation, results in the graphs are shown using a logarithmic scale, where higher logarithmic value represents minimum value of the mean. As shown in Figure 6.2, the proposed algorithm achieved a high logarithmic value of the mean (minimum value of the mean) for all metrics, indicating that the proposed algorithm is more efficient and robust in comparison to existing MOOAs.

The details related to a surgical patient (Table 6.8) were submitted as input to the STS module. Then, a suitable active patient cluster was selected from the six optimal active patient clusters (obtained from the SHM module), and associated surgical teams were extracted from it. The selected cluster contained 400 orthopaedic surgical records in which 15 distinct surgeons, 40 anesthesiologists, and 30 nurses were involved. It resulted in 18000 possible combinations of surgical teams. Finally, 135 surgical teams were generated as an initial search population, and the proposed algorithm was implemented on it. The results are shown in Figures 6.3 and 6.4. Figures 6.3(a) and 6.3(b) show the comparison between the proposed and existing algorithms in terms of the complication ratio and surgical feedback rating, respectively. From Figure 6.3(a) it is clear that the proposed algorithm converged faster to the optimal solution and obtained the lowest value of the complication ratio in comparison to existing algorithms. Similarly, Figure 6.3(b) illustrates that the proposed algorithm achieved maximum value of surgical feedback rating also. The final results presented in Figure 6.4 revealed that six optimal surgical teams were selected for the referred surgical patient. This can be assigned to the patient as per availability of the team members.

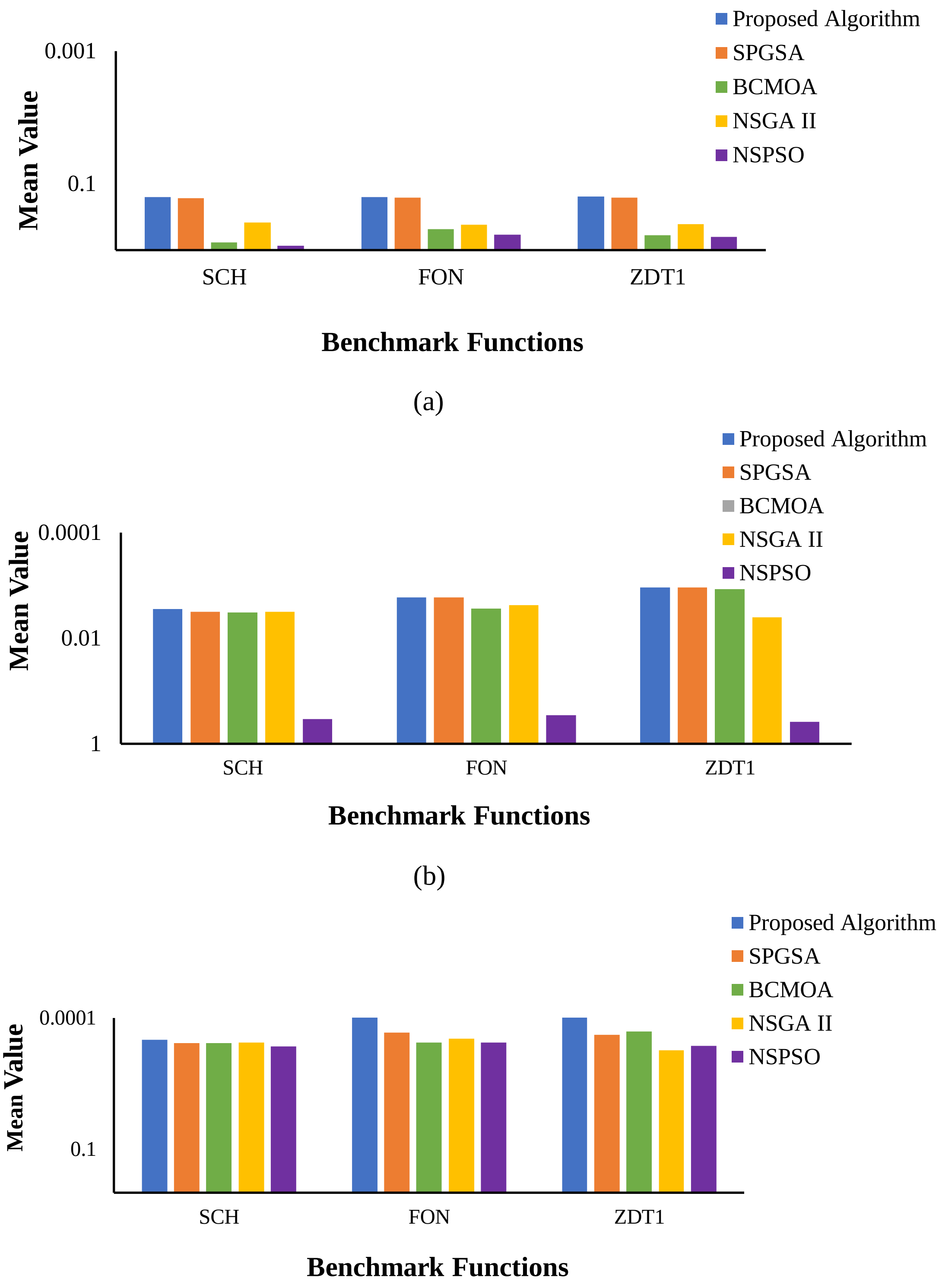
Table 6.6: Performance parameters used in the proposed multi-objective optimization algorithm

Description	**Parameter**	**Value for benchmark functions**	**Value for surgical team selection**
Population (Surgical team) size	P_{SSize}	100	135
External population size	P_{ESize}	100 for SCH, FON, and ZDT1	6
Initial value of Coulomb's constant	K_0	500	500
The maximum number of iterations	$MaxIT$	100 for SCH and FON, and 250 for ZDT1	50
Initial, Final crossover probability	P_{CR}, P_{CF}	1.0, 0.0	1.0, 0.0
Initial, Final mutation probability	P_{MI}, P_{MF}	0.01, 0.001	0.01, 0.001

Table 6.7: Comparison of the performance of proposed and existing multi-objective optimization algorithms

Performance Metric	**Benchmark Functions**		**Algorithms**				
			Proposed Algorithm	**SPGSA**	**NSGA II**	**NSPSO**	**BCMOA**
CM Metric	SCH	Mean	1.59×10^{-1}	1.65×10^{-1}	3.8×10^{-1}	8.6×10^{-1}	7.6×10^{-1}
	FON	Mean	1.58×10^{-1}	1.61×10^{-1}	4.14×10^{-1}	5.81×10^{-1}	4.8×10^{-1}
	ZDT1	Mean	1.56×10^{-1}	1.61×10^{-1}	4.06×10^{-1}	6.38×10^{-1}	5.9×10^{-1}
DM Metric	SCH	Mean	2.81×10^{-3}	3.2×10^{-3}	3.14×10^{-3}	3.40×10^{-1}	3.2×10^{-3}
	FON	Mean	1.68×10^{-3}	1.7×10^{-3}	2.36×10^{-3}	2.84×10^{-1}	2.7×10^{-3}
	ZDT1	Mean	1.09×10^{-3}	1.1×10^{-3}	4.02×10^{-3}	3.81×10^{-1}	1.1×10^{-3}
GD Metric	SCH	Mean	3.19×10^{-4}	3.78×10^{-4}	3.68×10^{-4}	4.5×10^{-4}	3.78×10^{-4}
	FON	Mean	3.65×10^{-5}	2.13×10^{-4}	2.94×10^{-4}	3.6×10^{-4}	3.62×10^{-4}
	ZDT1	Mean	3.02×10^{-5}	2.4×10^{-4}	5.56×10^{-4}	4.3×10^{-4}	2.02×10^{-4}

(~ shows results not available)



(a)

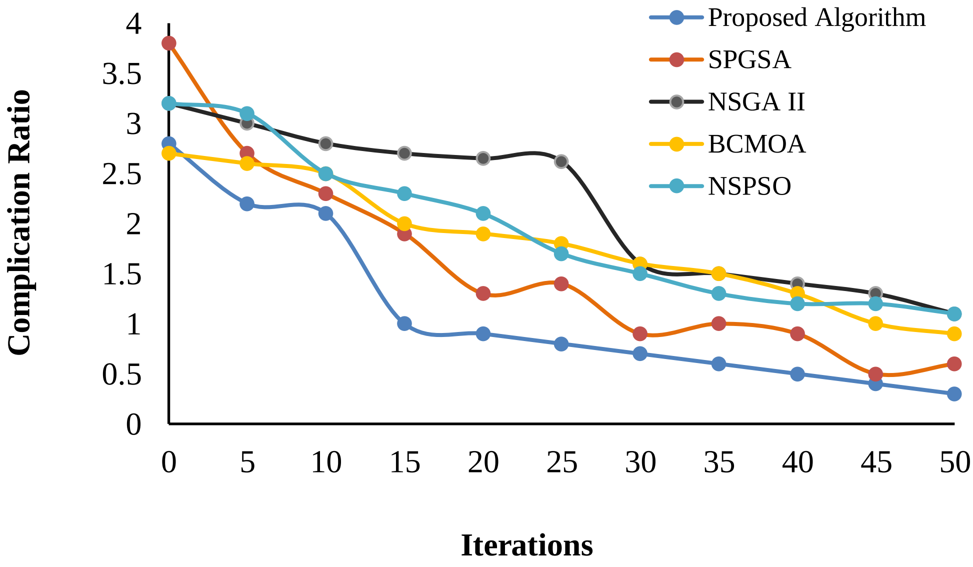
(b)

(c)

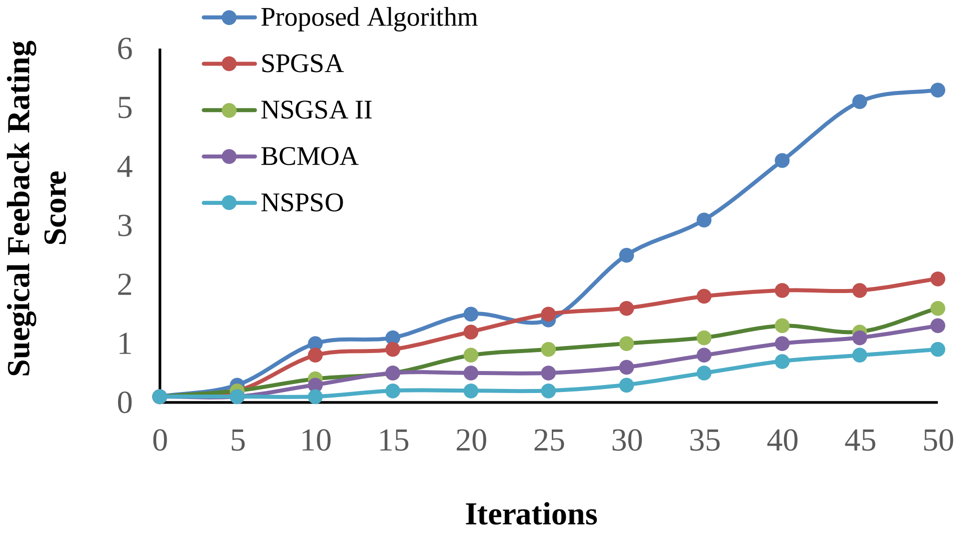
Figure 6.2: Comparison of performance between proposed and existing MOOAs based on a) CM, b) DM, and c) GD

Table 6.8: Attributes of a surgical patient

Attribute	Type	Description
Age	Numeric	62
Gender	Categorical	Male
BMI	Numeric	42
ASA fitness grade	Numeric	2
Marital Status	Categorical	Married
Ethnicity	Categorical	Indian
Comorbidity	Numeric	3
Type of Surgery	Categorical	Minor
Surgery Duration	Numeric	Length of surgical procedure
Procedural code	Categorical	0KQV0ZZ
Diagnose code	Categorical	S82.91XA
Surgery Domain	Categorical	Orthopaedic
Grade of Surgery	Categorical	Mild
Urgency of surgery	Categorical	Elective
LOS	Numeric	3.4 days



(a) Performance comparison based on complication ratio



(b) Performance comparison based on surgical feedback rating

Figure 6.3: Comparison of performance between proposed and existing MOOAs for surgical team selection based on; a) complication ratio, b) surgical feedback rating

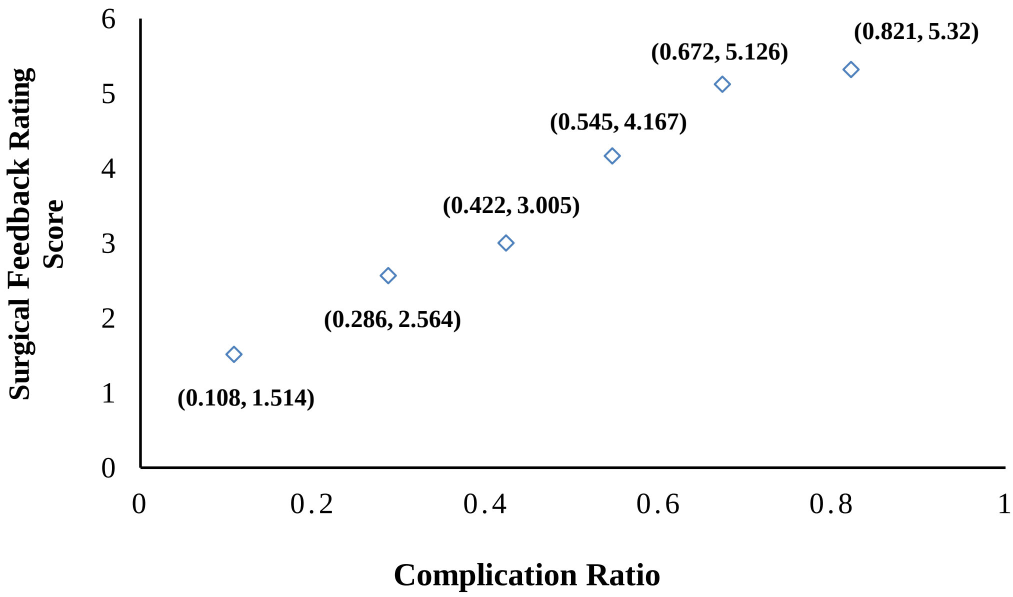


Figure 6.4: Optimal Pareto list of selected surgical team

Chapter 7

CONCLUSION AND FUTURE DIRECTIONS

7.1 Conclusion

In this thesis, the focus is on development of a framework for scheduling and optimization of healthcare resources. To accomplish this goal, we have proposed a framework containing three models for prioritizing patients from the waiting list, clustering surgical patients, and selecting optimal surgical teams using soft computing approaches. The first proposed model is a decision-making model PSWL-CCI, which aims to prioritize patients from the surgical waiting list. This model is comprised of four phases. In the first phase, factors (criteria and associated sub-criterion) related to patients are selected, and a hierarchical structure is formed. In the second phase, the decision-makers (experts) are asked to compare all factors and provide a rating for every comparison. Then, based on the unanimous decision (rating) of all experts, PCMs are generated. The third phase computes each PCM's consistency and validates it against the cosine consistency index (CCI) threshold value. The fourth phase is concerned with two individual activities: (i) each PCM that satisfies the CCI threshold in the third phase is used to compute priority of patients. The cosine maximization method (CM) with AHP is used to compute the resulting priority of patients. (ii) for PCMs, which fails to satisfy the CCI threshold, a hybrid algorithm HMWCA (Hybrid modified water cycle algorithm) is proposed to modify them to obtain PCMs with optimized CCI value. The proposed algorithm utilizes the features of a population-based meta-heuristic water cycle algorithm with evaporation-rate (ER-WCA), genetic algorithm, and 2-opt heuristic algorithm. In the proposed algorithm, the concepts of salt concentration and infiltration are introduced to the evaporation rate of ER-WCA, which extends ER-WCA to modified WCA. The proposed algorithm is tested on different inconsistent PCMs. Its performance is compared to existing algorithms on three criteria: improvement in CCI, optimal CCI score, and paired sample t-test. Experimental analysis reveal that the proposed algorithm outperforms the existing algorithms and generates statistically significant results. Finally, the proposed model is validated through a case study of the orthopedic surgery department at multispecialty Shri Mahant Indiresh Hospital in India. Data related to patients and hospital resources is acquired from doctors and other medical staff of the hospital. The performance of the proposed model is compared with existing four prioritization methods: EV, WLS, AN, and LLS using two parameters: Euclidean distance (ED) and minimum violation (MV). From the results it is observed that the proposed PSWL-CCI model achieves minimum ED=4.6 and MV=1, thus outperforms existing prioritization methods in terms of generating highly consistent and significantly better priorities.

Next, we have proposed an efficient clustering algorithm to enhance the hospital's existing surgical record management procedure (SRMP) to arrange surgical patients into optimal sub-groups. The proposed clustering algorithm is based on a population-based meta-heuristic, artificial electric field algorithm (AEFA), and is designed to deal with mixed dataset. The proposed algorithm requires no prior specification of the number of clusters. Instead, it identifies the number of clusters and cluster centers automatically. The proposed algorithm utilizes threshold setting and cut-off value for selecting and refining cluster centers. In the proposed algorithm, similarity between data points and different clusters is measured using a distance measure, i.e., Euclidean distance for numeric attributes and the probability of co-occurrence for categorical attributes. Firstly, the proposed algorithm is tested on five real-life datasets and compared with existing mixed data clustering algorithms based on two robustness measures: average accuracy and standard deviation. Results show that the proposed algorithm obtains higher average accuracy and lower standard deviation and outperforms in comparison to existing clustering algorithms. Then, the proposed algorithm is applied to a real postoperative surgical dataset obtained from a multispecialty hospital in India. Finally, the performance of the proposed clustering algorithm is validated using an unpaired t-test. Experimental results show that the proposed algorithm efficiently identifies the number of clusters and arranges surgical patients into 6 optimal sub-groups with an average accuracy of 0.87 and standard deviation of 0.12. It is also shown that the clustering results are robust and statistically significant.

Finally, we have addressed a vital challenge of multispecialty hospitals, where many surgical patients are treated, and it is crucial to arrange a suitable surgical team for them to obtain satisfactory surgical outcomes. We have proposed a decision-making model to select an optimal list of surgical teams for a referred surgical patient so that each patient can receive quality surgical care. The proposed model addresses two major issues of any multispecialty hospital: enhancement of existing surgical history management system and selection of an optimal surgical team. The proposed model is composed of two modules: the surgical history management (SHM) module and the surgical team selection (STS) module. SHM focuses on arranging the existing surgical patients into optimal patient sub-groups. It assists the STS module in selecting the optimal list of surgical teams. To identify the optimal sub-groups for the patients, the above proposed clustering algorithm is modified and used in the SHM module. For selection of an optimal list of the surgical teams, a population-based meta-heuristic algorithm AEFA for multi-objective optimization (MOOA) is proposed in the STS module. The complication ratio and surgical feedback rating are used in the proposed MOOA to evaluate performance of the surgical team. The performance of the proposed clustering algorithm is compared with existing non-mixed dataset clustering algorithms using four real-life datasets on the basis of three cluster quality measures: SI index, DI index, and DB index. Then, the performance of the proposed algorithm is compared with existing mixed dataset clustering algorithm using five real-life datasets on two performance measures: average accuracy and standard deviation. The results show that the proposed clustering algorithm performs better in comparison to existing clustering algorithms. Subsequently, the proposed clustering algorithm is also validated through a case study of the orthopaedic surgery department as discussed earlier. Results demonstrate that the proposed clustering algorithm identifies 6 optimal clusters of patients with the highest average accuracy of 0.97 and lowest standard deviation of 0.13. Similarly, the performance of the proposed MOOA is compared with the existing MOOAs using three benchmark functions: SCH, FON, and ZDT1. Results show that the proposed MOOA produces better results in comparison to the existing MOOAs. Finally, the proposed MOOA is applied to the surgical team dataset, and its performance is compared with existing MOOAs based on complication ratio and surgical feedback rating. The results show that the proposed MOOA achieves better and optimal surgical teams in comparison to the existing MOOA in terms of lower complication ratio and higher surgical feed rating results. The overall results demonstrate that the proposed model shows its efficacy not only in clustering surgical patients into optimal subgroup but also in selecting the optimal list of surgical teams.

7.2 Future Directions

The research work described in the thesis has a number of promising directions for future research. In the future, the inconsistent pair-wise comparison matrix (PCM), which is improved using HMWCA, can be further improved for more optimal outcomes. The proposed algorithm can be extended to handle incomplete, inconsistent PCMs also. The scope of proposed patient prioritization model is limited to prioritize elective surgical waiting list. The proposed model can be extended to help hospitals to decide during emergency or critical care services.

The proposed clustering algorithm can be ensemble with other meta-heuristic approaches for solving complex high dimensional clustering problems. Further, this algorithm can be modified to improve the optimal selection of clusters.

In the proposed optimal surgical team selection model, the proposed surgical team selection module can be enhanced for more performance measures. Subsequently, the proposed multi-objective optimization algorithm, can be integrated with other meta-heuristic approaches to improve performance and achieve better optimization.

References

[1] S. Shalev, “Solitary Confinement As a Prison Health Issue,” in *Prisons and Health*, S. Enggist, L. Moller, G. Galea, and C. Udesen, Eds., Copenhagen, Denmark: World Health Organization, 2017, pp. 27-35.

[2] M. P. Kieny, T. G. Evans, S. Scarpetta, E. T. Kelley, N. Klazinga, I. Forde, J. H. M. Veillard, S. Leatherman, S. Syed, S. M. Kim, and S. B. Nejad., *Delivering quality health services: a global imperative for universal health coverage*, Washington, DC: World Bank Group, 2018, pp. 1-100.

[3] L. Luxon, “Infrastructure–the key to healthcare improvement,” *Future Hospital Journal*, vol. 2, no.1, pp. 4-7, 2015.

[4] S. Batun, and M. A. Begen, “Optimization in healthcare delivery modeling: Methods and applications,” in *Handbook of Healthcare Operations Management*, Springer, New York, 2013, pp. 75-119.

[5] A. M. Mosadeghrad, “Factors influencing healthcare service quality,” *International Journal of Health Policy and Management*, vol. 3, no. 2, pp. 77-89, 2014.

[6] H. Ransom, and J. M. Olsson, “Allocation of Health Care Resources: Principles for Decision-making,” *Pediatrics in Review*, vol. 38, no. 7, pp. 320-329, 2017.

[7] World Health Organization, ‘‘The world health report 2006: working together for health,” World Health Organization, Geneva, 2006.

[8] B. I. Helgheim, B. E. Sandbaek, and L. Slyngstad, “A prospective investigation of direct and indirect home care activities in three rural Norwegian municipalities,” *BMC Health Services Research*, vol. 18, no. 1, pp. 1-8, 2018.

[9] L. E. Kakushi, and Y. D. M. Évora, “Direct and indirect nursing care time in an intensive care unit,” *Revista Latino-Americana De Enfermagem*, vol. 22, no. 1, pp. 150-157, 2014.

[10] P. K. Panigrahi, “A comparative study of supervised machine learning techniques for spam e-mail filtering,” in *International Conference on Computational Intelligence and Communication Networks*, pp. 506-512, IEEE, Mathura, India, 2012.

[11] F. Matsebula, and E. Mnkandla, “A big data architecture for learning analytics in higher education,” in *IEEE AFRICON*, pp. 951-956, IEEE, Cape Town, South Africa, 2017.

[12] M. Mahdavi, M. Parsaeian, E. Jaafaripooyan, and S. Ghaffari, “Recent Iranian health system reform: an operational perspective to improve health services quality,” *International Journal of Health Policy and Management*, vol. 7, no. 1, pp. 70-74, 2018.

[13] M. M. Malik, S. Abdallah, and M. Ala’raj, “Data mining and predictive analytics applications for the delivery of healthcare services: a systematic literature review,” *Annals of Operations Research*, vol. 270, no. 1, pp. 287-312, 2018.

[14] P. E. Liss, “Allocation of scarce resources in health care: Values and concepts,” *Texto & Contexto-Enfermagem*, vol. 15, no. SPE, pp. 125-134, 2006.

[15] Q. Ye, Z. Deng, Y. Chen, J. Liao, G. Li, and Y. Lu, “How Resource Scarcity and Accessibility Affect Patients’ Usage of Mobile Health in China: Resource Competition Perspective,” *JMIR mHealth and uHealth*, vol. 7, no. 8, pp. e13491, 2019.

[16] L. V. Green, J. Soares, J. F. Giglio, and R. A. Green, “Using Queueing Theory to Increase the Effectiveness of Emergency Department Provider Staffing,” *Academic Emergency Medicine*, vol. 13, no. 1, pp. 61-68, 2006.

[17] C. W. Chan, V. F. Farias, N. Bambos, and G. J. Escobar, “Optimizing Intensive Care Unit Discharge Decisions with Patient Readmissions,” *Operations Research*, vol. 60, no. 6, pp. 1323-1341, 2012.

[18] S. Gambhir, S. K. Malik, and Y. Kumar, “Role of soft computing approaches in healthcare domain: a mini review,” *Journal of Medical Systems*, vol. 40, no. 12, pp. 1-20, 2016.

[19] A. Bhatia, V. Mago, and R. Singh, “Use of soft computing techniques in medical decision making: A survey,” in *International Conference on Advances in Computing, Communications and Informatics*, pp. 1131-1137, IEEE, Delhi, India, 2014.

[20] D. R. Anderson, “The impact of resource management on hospital efficiency and quality of care,” Ph.D. dissertation, University of Maryland, College Park, MD, USA, 2013.

[21] A. Testi, E. Tanfani, R. Valente, G. L. Ansaldo, and G. C. Torre, “Prioritizing surgical waiting lists,” *Journal of Evaluation in Clinical Practice*, vol. 14, no. 1, pp. 59-64, 2008.

[22] G. Mariotti, L. Siciliani, V. Rebba, R. Fellini, M. Gentilini, G. Benea, P. Bertoli, L. Bistolfi, S. Brugaletta, P. Camboa, and P. Casucci, “Waiting time prioritisation for specialist services in Italy: the homogeneous waiting time groups approach,” *Health Policy*, vol. 117, no. 1, pp. 54-63, 2014.

[23] S. A. Rahimi, and A. Jamshidi, “Prioritization of organ transplant patients using analytic network process,” in *Proceedings of Industrial and Systems Engineering Research Conference*, pp. 3544-3551, Montreal, Canada, 2014.

[24] S. A. Rahimi, A. Jamshidi, A. Ruiz, and D. Aït-Kadi, “A new dynamic integrated framework for surgical patients' prioritization considering risks and uncertainties,” *Decision Support Systems,* vol. 88, pp. 112-120, 2016.

[25] K. Kaur, and R. Rani, “Multi-Model Persistent Solution for Healthcare Big Data,” *Journal of Medical Imaging and Health Informatics*, vol. 6. no. 4, pp. 937-947, 2016.

[26] L. Mikhailov, “A fuzzy programming method for deriving priorities in the analytic hierarchy process,” *Journal of the Operational Research Society*, vol. 51, no. 3, pp. 341-349, 2000.

[27] C. S. Lin, and S. L. Harris, “A unified framework for the prioritization of organ transplant patients: Analytic hierarchy process, sensitivity and multifactor robustness study,” *Journal of Multi-Criteria Decision Analysis*, vol. 20, no. 3-4, pp. 157-172, 2013.

[28] S. A. Rahimi, A. Jamshidi, A. Ruiz, and D. Ait-Kadi, “Multi-criteria decision making approaches to prioritize surgical patients,” *in Proceedings of Health Care Systems Engineering for Scientists and Practitioners,* pp. 25-34, Springer, Cham, 2016.

[29] I. E. Evangelou, D. G. Hadjimitsis, A. A. Lazakidou, and C. Clayton, “Data mining and knowledge discovery in complex image data using artificial neural networks,” in *Proceedings of Workshop on Complex Reasoning on Geographical Data*, Paphos, Cyprus, 2001.

[30] B. S. Everitt. *Cluster Analysis*, 3rd ed. London, England: Edward Arnold / Halsted Press, 1993.

[31] M. R. Rao, "Cluster analysis and mathematical programming," *Journal of American Statistical Association*, vol. 66, no. 335, pp. 622–626, 1971.

[32] A. J. Graaff, and A. P. Engelbrecht, "Using sequential deviation to dynamically determine the number of clusters found by a local network neighbourhood artificial immune system," *Applied Soft Computing*, vol. 11 no. 2, pp. 2698–2713, 2011.

[33] H. Frigui, and R. Krishnapuram, "A robust competitive clustering algorithm with applications in computer vision," *IEEE Transections on Pattern Analysis and Machine Intelligence*, vol. 21, no. 5, pp. 450–465, 1999.

[34] Y. Leung, J. S. Zhang, and Z. B. Xu, "Clustering by scale-space filtering," *IEEE Transections on Pattern Analysis and Machine Intelligence*, vol. 22, no. 12, pp. 1396-1410, 2000.

[35] Y. Liu, X. Wu, and Y. Shen, "Automatic clustering using genetic algorithms," *Applied Mathematics and Computation*, vol. 218, no. 4, pp. 1267-1279, 2011.

[36] Z. Aliniya, and S. A. Mirroshandel, "A novel combinatorial merge-split approach for automatic clustering using imperialist competitive algorithm," *Expert Systems with Applications*, vol. 117, no. 1, pp. 243–266, 2019.

[37] H. He, and Y. Tan, "A two-stage genetic algorithm for automatic clustering," *Neurocomputing*, vol. 81, no. 1, pp. 49-59, 2012.

[38] D. Doval, S. Mancoridis, and B. S. Mitchell, "Automatic clustering of software systems using a genetic algorithm," in *Proceedings of 9th International Workshop Software Technology and Engineering Practice*, pp. 73–81, IEEE, Pittsburgh, PA, USA 1999.

[39] Z. Izakian, M. S. Mesgari, and A. Abraham, “Automated clustering of trajectory data using a particle swarm optimization,” *Computers, Environment and Urban Systems*, vol. 55, pp. 55-65, 2016.

[40] S. Das, A. Abraham, and A. Konar, “Automatic kernel clustering with a multi-elitist particle swarm optimization algorithm,” *Pattern Recognition Letters*, vol. 29, no. 5, pp. 688-699, 2008.

[41] V. Kumar, J. K. Chhabra, and D. Kumar, “Automatic cluster evolution using gravitational search algorithm and its application on image segmentation,” *Engineering Applications of Artificial Intelligence*, vol. 29, pp. 93–103, 2014.

[42] V. Kumar, and D. Kumar, “Automatic clustering and feature selection using gravitational search algorithm and its application to microarray data analysis,” *Neural Computing and Applications*, vol. 31, no. 8, pp. 3647-3663, 2019.

[43] S. Das, A. Abraham, and A. Konar, “Automatic clustering using an improved differential evolution algorithm,” *IEEE Transactions on systems, man, and cybernetics-Part A: Systems and Humans*, vol. 38, no. 1, pp. 218–237, 2007.

[44] S. Das, A. Chowdhury, and A. Abraham, “A bacterial evolutionary algorithm for automatic data clustering,” in *IEEE Congress on Evolutionary Computation*, pp. 2403–2410, Trondheim, Norway, 2009.

[45] R. J. Kuo, Y. D. Huang, C. C. Lin, Y. H. Wu, and F. E. Zulvia, “Automatic kernel clustering with bee colony optimization algorithm,” *Information Sciences*, vol. 283, pp. 107–122, 2014.

[46] Z. G. Su, P. H. Wang, J. Shen, Y. G. Li, Y. F. Zhang, and E.J. Hu, “Automatic fuzzy partitioning approach using Variable string length Artificial Bee Colony (VABC) algorithm,” *Applied Soft Computing*, vol. 12, no. 11, pp. 3421–3441, 2012.

[47] A. José-García, and W. Gómez-Flores, "Automatic clustering using nature-inspired metaheuristics: a survey," *Applied Soft Computing*, vol. 41, pp. 192–213, 2016.

[48] O. E. Abbasoglu, E. C. Sener, and A. S. Sanac, "Factors influencing the successful outcome and response in strabismus surgery," *Eye*, vol. 10, no. 3, pp. 315-320, 1996.

[49] A. Jones, "Multidisciplinary team working: Collaboration and conflict," *International Journal of Mental Health Nursing*, vol. 15, no. 1, pp. 19-28, 2006.

[50] R. Guimera, B. Uzzi, J. Spiro, and L. A. N Amaral, "Team assembly mechanisms determine collaboration network structure and team performance," *Science*, vol. 308, no. 5722, pp. 697-702, 2005.

[51] M. A. Makary, J. B. Sexton, J. A. Freischlag, E. A. Millman, D. Pryor, C. Holzmueller, and P. J. Pronovost, "Patient safety in surgery," *Annals of surgery*, vol. 243, no. 5, pp. 628-635, 2006.

[52] J. B. Sexton, M. A. Makary, A. R. Tersigni, D. Pryor, A. Hendrich, E. J. Thomas, C. G. Holzmueller, A. P. Knight, Y. Wu, and P. J. Pronovost, "Teamwork in the operating room frontline perspectives among hospitals and operating room personnel," *Anesthesiology: The Journal of the American Society of Anesthesiologists,*vol. 105, no. 5, pp. 877-884, 2006.

[53] Y. Y. Hu, A. F. Arriaga, E. M. Roth, S. E. Peyre, K. A. Corso, R. S. Swanson, R. T. Osteen, P. Schmitt, A. M. Bader, M. J. Zinner, and C. C. Greenberg, "Protecting patients from an unsafe system: the etiology & recovery of intra-operative deviations in care," *Annals of Surgery*, vol. 256, no. 2, pp. 203-210, 2012.

[54] K. Mazzocco, D. B. Petitti, K. T. Fong, D. Bonacum, J. Brookey, S. Graham, R. E. Lasky, J. B. Sexton, and E. J. Thomas, "Surgical team behaviors and patient outcomes," *The American Journal of Surgery*, vol. 197, no. 5, pp. 678-685, 2009.

[55] A. Ebadi, P. J. Tighe, L. Zhang, and P. Rashidi, "DisTeam: A decision support tool for surgical team selection," *Artificial Intelligence in Medicine*, vol. 76, pp. 16-26, 2017.

[56] J. Carthey, M. R. de Leval, and J. T. Reason, "The human factor in cardiac surgery: errors and near misses in a high technology medical domain," *The Annals of Thoracic Surgery*, vol. 72, no. 1, pp. 300-305, 2001.

[57] S. E. Regenbogen, C. C. Greenberg, D. M. Studdert, S. R. Lipsitz, M. J. Zinner, and A. A. Gawande, "Patterns of technical error among surgical malpractice claims: an analysis of strategies to prevent injury to surgical patients," *Annals of Surgery*, vol. 246, no. 5, pp. 705-711, 2007.

[58] Jr S. O. Rogers, A. A. Gawande, M. Kwaan, A. L. Puopolo, C. Yoon, T. A. Brennan, and D. M. Studdert "Analysis of surgical errors in closed malpractice claims at 4 liability insurers," *Surgery*, vol. 140, no. 1, pp. 25-33, 2006.

[59] L. L. Faltz, J. N Morley, E. Flink, and P. D. Dameron, "The New York Model: Root Cause Analysis Driving Patient Safety Initiative to Ensure Correct Surgical and Invasive Procedures," in *Advances in Patient Safety: New Directions and Alternative Approaches*, vol.1, K. Henriksen, J. B. Battles, M. A. Keyes, and M. L. Grady, Eds., Rockville, MD, USA: Agency for Healthcare Research and Quality, 2008.

[60] R. Mallett, M. Conroy, L. Z. Saslaw, and S. Moffatt-Bruce, "Preventing wrong site, procedure, and patient events using a common cause analysis," *American Journal of Medical Quality*, vol. 27, no. 1, pp. 21-29, 2012.

[61] K. Catchpole, A. Mishra, A. Handa, and P. McCulloch, "Teamwork and error in the operating room: analysis of skills and roles," *Annals of Surgery*, vol. 247, no. 4, pp. 699-706, 2008.

[62] C. K. Christian, M. L. Gustafson, E. M. Roth, T. B. Sheridan, T. K. Gandhi, K. Dwyer, M. J. Zinner, and M. M. Dierks, "A prospective study of patient safety in the operating room," *Surgery*, vol. 139, no. 2, pp. 159-173, 2006.

[63] L. Lingard, R. Reznick, S. Espin, G. Regehr, and I. DeVito, "Team communications in the operating room: talk patterns, sites of tension, and implications for novices," *Academic Medicine*, vol. 77, no. 3, pp. 232-237, 2002.

[64] S. Guerlain, R. B. Adams, F. B. Turrentine, T. Shin, H. Guo, S. R. Collins, and J. F. Calland, "Assessing team performance in the operating room: development and use of a "black-box" recorder and other tools for the intraoperative environment," *Journal of the American College of Surgeons*, vol. 200, no. 1, pp. 29-37, 2005.

[65] M. Michaelson, and L. Levi, "Videotaping in the admitting area: a most useful tool for quality improvement of the trauma care," *European Journal of Emergency Medicine: Official Journal of the European Society for Emergency Medicine*, vol. 4, no. 2, pp. 94-96, 1997.

[66] S. K. Howard, D. M. Gaba, K. J. Fish, G. Yang, and F. H. Sarnquist, "Anesthesia crisis resource management training: teaching anesthesiologists to handle critical incidents," *Aviation, Space, and Environmental Medicine*, vol. 63, no. 9, pp. 763-770, 1992.

[67] L. S. Leach, R. C. Myrtle, F. A. Weaver, and S. Dasu, "Assessing the performance of surgical teams," *Health Care Management Review*, vol. 34, no. 1, pp. 29-41, 2009.

[68] L. S. Franz, and J. L. Miller, "Scheduling medical residents to rotations: solving the large-scale multiperiod staff assignment problem," *Operations Research*, vol. 41, no. 2, pp. 269-279, 1993.

[69] D. C. Hador, and Steering Committee of the Western Canada Waiting List Project, "Setting priorities for waiting lists: defining our terms," *Canadian Medical Association Journal*, vol. 163, no. 7, pp. 857-860, 2000.

[70] M. M. Dowsey, J. Gunn, and P. F. Choong, "Selecting those to refer for joint replacement: who will likely benefit and who will not?," *Best Practice & Research Clinical Rheumatology*, vol. 28, no. 1, pp. 157-171, 2014.

[71] A. Lack, R. T. Edwards, and A. Boland, "Weights for waits: lessons from Salisbury," *Journal of Health Services Research & Policy*, vol. 5, no. 2, pp. 83-88, 2000.

[72] M. Solans-Domènech, P. Adam, C. Tebé, and M. Espallargues, "Developing a universal tool for the prioritization of patients waiting for elective surgery," *Health Policy*, vol. 113, no. 1-2, pp. 118-126, 2013.

[73] R. Valente, A. Testi, E. Tanfani, M. Fato, I. Porro, M. Santo, G. Santori, G. Torre, and G. Ansaldo, "A model to prioritize access to elective surgery on the basis of clinical urgency and waiting time," *BMC Health Services Research*, vol. 9, no. 1, pp. 1-15, 2009.

[74] M. Ahlert, and H. Kliemt, "Problems of priority change in kidney allocation and beyond," *The European Journal of Health Economics*, vol. 14, no. 3, pp. 383-390, 2013.

[75] L. Al-Ebbini, A. Oztekin, and Y. Chen, "FLAS: Fuzzy lung allocation system for US-based transplantations," *European Journal of Operational Research*, vol. 248, no. 3, pp. 1051-1065, 2016.

[76] E. Cholongitas, G. Germani, and A. K. Burroughs, “Prioritization for liver transplantation,” *Nature Reviews Gastroenterology & Hepatology*, vol. 7, no. 12, pp. 659-668, 2010.

[77] D. Min, and Y. Yih, “An elective surgery scheduling problem considering patient priority” *Computers & Operations Research*, vol. 37, no. 6, pp. 1091-1099, 2010.

[78] E. B. Fields, G. E. Okudan, and O. M. Ashour, “Rank aggregation methods comparison: A case for triage prioritization,” *Expert Systems with Applications*, vol. 40, no. 4, pp. 1305-1311, 2013.

[79] M. Yang, M. J. Fry, J. Raikhelkar, C. Chin, A. Anyanwu, J. Brand, and C. Scurlock, “A model to create an efficient and equitable admission policy for patients arriving to the cardiothoracic ICU,” *Critical Care Medicine*, vol. 41, no. 2, pp. 414-422, 2013.

[80] A. T. W. Chu, R. E. Kalaba, and K. Spingarn, “A comparison of two methods for determining the weights of belonging to fuzzy sets,” *Journal of Optimization Theory and Applications*, vol. 27, no. 4, pp. 531-538, 1979.

[81] T. L. Saaty, “Eigenvector and logarithmic least squares,” *European Journal of Operational Research*, vol. 48, no. 1, pp. 156-160, 1990.

[82] B. Srdjevic, “Combining different prioritization methods in the analytic hierarchy process synthesis,” *Computers & Operations Research*, vol. 32, no. 7, pp. 1897-1919, 2005.

[83] G. Crawford, and C. Williams, “A note on the analysis of subjective judgment matrices,” *Journal of Mathematical Psychology*,” vol. 29, no. 4, pp. 387-405, 1985.

[84] K. O. Cogger, and P. L. Yu, “Eigenweight vectors and least-distance approximation for revealed preference in pairwise weight ratios,” *Journal of Optimization Theory and Applications*, vol. 46, no. 4, pp. 483-491, 1985.

[85] L. Mikhailov, and M. G. Singh, “Comparison analysis of methods for deriving priorities in the analytic hierarchy process,” *in Proceedings of IEEE International Conference on Systems, Man, and Cybernetics*, vol. 1, pp. 1037-1042, Tokyo, Japan, 1999.

[86] C. Lin, G. Kou, and D. Ergu, “A statistical approach to measure the consistency level of the pairwise comparison matrix,” *Journal of the Operational Research Society*, vol. 65, no. 9, pp. 1380-1386, 2014.

[87] S. Siraj, L. Mikhailov, and J. Keane, “A heuristic method to rectify intransitive judgments in pairwise comparison matrices,” *European Journal of Operational Research*, vol. 216, no. 2, pp. 420-428, 2012.

[88] Z. Wu, and J. Xu, “A consistency and consensus based decision support model for group decision making with multiplicative preference relations,” *Decision Support Systems*, vol. 52, no. 3, pp. 757-767, 2012.

[89] D. Cao, L. C. Leung, and J. S. Law, “Modifying inconsistent comparison matrix in analytic hierarchy process: A heuristic approach,” *Decision Support Systems*, vol. 44, no. 4, pp. 944-953, 2008.

[90] G. Khatwani, and A. K. Kar, “Improving the Cosine Consistency Index for the analytic hierarchy process for solving multi-criteria decision making problems,” *Applied Computing and Informatics*, vol. 13, no. 2, pp. 118-129, 2017.

[91] A. S. Girsang, C. W. Tsai, and C. S. Yang, “Ant algorithm for modifying an inconsistent pairwise weighting matrix in an analytic hierarchy process,” *Neural Computing and Applications*, vol. 26, no. 2, pp. 313-327, 2015.

[92] D. Ergu, G. Kou, Y. Peng, and Y. Shi, “A simple method to improve the consistency ratio of the pair-wise comparison matrix in ANP,” *European Journal of Operational Research*, vol. 213, no. 1, pp. 246-259, 2011.

[93] J. F. da Serra Costa, “A genetic algorithm to obtain consistency in analytic hierarchy process,” *Brazilian Journal of Operations & Production Management*, vol. 8, no. 1, pp. 55-64, 2011.

[94] C. C. Lin, W. C. Wang, and W. D. Yu, “Improving AHP for construction with an adaptive AHP approach (A3),” *Automation in Construction*, vol. 17, no. 2, pp. 180-187, 2008.

[95] S. Kaur, H. Aggarwal, and R. Rani, “Diagnosis of Parkinson's Disease Using Principle Component Analysis and Deep Learning,” *Journal of Medical Imaging and Health Informatics*, vol. 9, no. 3, pp. 602-609, 2019.

[96] A. Sharma, and R. Rani, “An optimized framework for cancer classification using deep learning and genetic algorithm,” *Journal of Medical Imaging and Health Informatics*, vol. 7, no. 8, pp. 1851-1856, 2017.

[97] P. M. Kumar, and R. P. Kumar, “Enhancing Bio-Medical Mammography Image Fusion Using Optimized Genetic Algorithm,” *Journal of Medical Imaging and Health Informatics*, vol. 9, no. 3, pp. 502-507, 2019.

[98] I. T. Yang, W. C. Wang, and T. I. Yang, “Automatic repair of inconsistent pairwise weighting matrices in analytic hierarchy process,” *Automation in Construction*, vol. 22, pp. 290-297, 2012.

[99] A. Sadollah, H. Eskandar, A. Bahreininejad, and J. H. Kim, “Water cycle algorithm with evaporation rate for solving constrained and unconstrained optimization problems,” *Applied Soft Computing*, vol. 30, pp. 58-71, 2015.

[100] E. Osaba, J. Del Ser, A. Sadollah, M. N. Bilbao, and D. Camacho, “A discrete water cycle algorithm for solving the symmetric and asymmetric traveling salesman problem,” *Applied Soft Computing*, vol. 71, pp. 277-290, 2018.

[101] S. Qiao, Y. Zhou, Y. Zhou, and R. Wang, “A simple water cycle algorithm with percolation operator for clustering analysis,” *Soft Computing*, vol. 23, no. 12, pp. 4081-4095, 2019.

[102] K. Gao, Y. Zhang, A. Sadollah, A. Lentzakis, and R. Su, “Jaya, harmony search and water cycle algorithms for solving large-scale real-life urban traffic light scheduling problem,” *Swarm and Evolutionary Computation*, vol. 37, pp. 58-72, 2017.

[103] G. Kou, and C. Lin, “A cosine maximization method for the priority vector derivation in AHP,” *European Journal of Operational Research*, vol. 235, no. 1, pp. 225-232, 2014.

[104] A. K. Jain and R.C. Dubes, *Algorithms for Clustering Data*. Upper Saddle River, New Jersey, USA: Prentice-Hall, Inc., 1988.

[105] J. MacQueen, “Some methods for classification and analysis of multivariate observations,” *in Proceedings of the 5th Berkeley Symposium on Mathematical Statistics and Probability*, vol. 1, no. 14, pp. 281–297, 1967.

[106] A. Mishra, and P. K. Mishra, “A randomized scheduling algorithm for multiprocessor environments using local search,” *Parallel Processing Letters*, vol. 26, no. 01, pp. 1650002 (1-17), 2016.

[107] R. Z. Khan, and M. H. K. Jabbar, “Improvement of Neuro Fuzzy Systems Using K-means Clustering,” *International Journal of Emerging Engineering research and Technology*, vol. 1, no. 2, pp. 1-7, 2013.

[108] J. S. Zhang, and Y. W. Leung, "Improved possibilistic c-means clustering algorithms," *IEEE Transactions on Fuzzy Systems*, vol. 12, no. 2, pp. 209-217, 2004.

[109] P. Berkhin, "A survey of clustering data mining techniques," in *Grouping Multidimensional Data*, Berlin, Heidelberg, Germany: Springer, 2006, pp. 25–71.

[110] T. Velmurugan, and T. Santhanam, "A survey of partition based clustering algorithms in data mining: an experimental approach," *Information Technology Journal*, vol. 10, no. 3, pp. 478-484, 2011.

[111] R. J. Kuo, H. S. Wang, T. L. Hu, and S. H. Chou, "Application of ant K-means on clustering analysis," *Computers & Mathematics with Applications* vol. 50, no. 10-12, pp. 1709–1724, 2005.

[112] A. Sami, and K. Tommi, "Introduction to Partitioning-based Clustering Methods with a Robust Example," Department of Mathematical Information Technology, Series C, Software Engineering and Computational Engineering 1/2006, University of Jyväskylä, Finland, 2006.

[113] M. Sarkar, B. Yegnanarayana, and D. Khemani, "A clustering algorithm using an evolutionary programming-based approach," *Pattern Recognition Letters,* vol. 18, no. 10, pp. 975-986, 1997.

[114] Y. Chen, C. Tang, J. Zhu, C. Li, S. Qiao, R. Li, and J. Wu, "Clustering without prior knowledge based on gene expression programming," in *3rd IEEE International Conference on Natural Computation,* pp. 451–455, Haikou, China, 2007.

[115] J. Liu, Y. Chi, Z. Liu, and S. He, "Ensemble multi-objective evolutionary algorithm for gene regulatory network reconstruction based on fuzzy cognitive

maps,” *CAAI Transections on Intelligence Technology*, vol. 4, no. 1, pp. 24–36, 2019.

[116] C. Y. Lee, and E. K. Antonsson, “Dynamic partitional clustering using evolution strategies”, in *IEEE International Conference on Industrial Electronics, Control and Instrumentation*, vol. 4, pp. 2716–2721, Nagoya, Japan, 2000.

[117] L. Y. Tseng, and S. B. Yang, “A genetic approach to the automatic clustering problem,” *Pattern Recognition*, vol. 34, no. 2, pp. 415–424, 2001.

[118] S. Bandyopadhyay, and U. Maulik, “Nonparametric genetic clustering: comparison of validity indices,” *IEEE Transections on System, Man, and Cybernetics, Part C (Application and Reviews)*, vol. 31, no. 1, pp. 120–125, 2001.

[119] S. Bandyopadhyay, and S. Saha, "A point symmetry-based clustering technique for automatic evolution of clusters,” *IEEE Transection on Knowledge and Data Engineering*, vol. 20, no. 11, pp. 1441–1457, 2008.

[120] A. Ahmad, and L. Dey, “A K-mean clustering algorithm for mixed numeric and categorical data,” *Data & Knowledge Engineering*, vol. 63, no. 2, pp. 503–527, 2007.

[121] A. Ahmad, and S. Hashmi, “K-Harmonic means type clustering algorithm for mixed datasets,” *Applied Soft Computing*, vol. 48, pp. 39–49, 2016.

[122] D. X. Chang, X. D. Zhang, C. W. Zheng, and D. M. Zhang, “A robust dynamic niching genetic algorithm with niche migration for automatic clustering problem,” *Pattern Recognition*, vol. 43, no. 4, pp. 1346–1360, 2010.

[123] S. M. Pan, and K. S. Cheng, “Evolution-based tabu search approach to automatic clustering,” *IEEE Transections on System, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 37, no. 5, pp. 827–838, 2007.

[124] M. G. Omran, A. Salman, and A. P. Engelbrecht, “Dynamic clustering using particle swarm optimization with application in image segmentation,” *Pattern Analysis and Applications*, vol. 8, no. 4, pp. 332-344, 2006.

[125] A. Chaudhary, and V. Bhattacharjee, “An efficient method for brain tumor detection and categorization using MRI images by K-means clustering & DWT,” *International Journal of Information Technology*, vol. 12, no. 1, pp. 141-148, 2020.

[126] A. Almagbile, “Estimation of crowd density from UAVs images based on corner detection procedures and clustering analysis,” *Geo-spatial Information Science*, vol. 22, no. 1, pp. 23–34, 2019.

[127] T. Cura, “A particle swarm optimization approach to clustering,” *Expert System with Applications*, vol. 39, no. 1, pp. 1582–1588, 2012.

[128] A. Chowdhury, S. Bose, and S. Das, “Automatic clustering based on invasive weed optimization algorithm,” in *International Conference on Swarm, Evolutionary, and Memetic Computing*, pp. 105–112, Springer, 2011.

[129] M. J. Sewitch, K. Leffondre, and P. L. Dobkin, “Clustering patients according to health perceptions: relationships to psychosocial characteristics and medication nonadherence,” *Journal of Psychosomatic Research*, vol. 56, no. 3, pp. 323-332, 2004.

[130] P. Finamore, M. A. Spruit, and J. M. Schols, R. A. Incalzi, E. F. Wouters and D. J. Janssen, “Clustering of patients with end-stage chronic diseases by symptoms: a new approach to identify health needs,” *Aging Clinical and Experimental Research*, vol. 33, no. 2, pp. 1-11, 2020.

[131] D. Gamberger, B. Ženko, A. Mitelpunkt, N. Shachar, and N. Lavrač, “Clusters of male and female Alzheimer’s disease patients in the Alzheimer’s Disease

Neuroimaging Initiative (ADNI) database," *Brain Informatics*, vol. 3, no. 3, pp. 169-179, 2016.

[132] H. Petwal, and R. Rani, "Prioritizing the Surgical Waiting List-Cosine Consistency Index: An Optimized Framework for Prioritizing Surgical Waiting List," *Journal of Medical Imaging and Health Informatics*, vol. 10, no. 12, pp. 2876-2892, 2020.

[133] S. R. Newcomer, J. F. Steiner, and E. A. Bayliss, "Identifying subgroups of complex patients with cluster analysis," *The American Journal of Managed Care*, vol. 17, no. 8, pp. e324-32, 2011.

[134] Y. Li, B. Rosenfeld, H. Pessin, and W. Breitbart, "Bayesian Nonparametric Clustering of Patients with Advanced Cancer on Anxiety and Depression," in *IEEE International Conference on Machine Learning and Applications*, pp. 674-678, Cancun, Maxico, 2017.

[135] C. Yu, B. T. Baune, K. A. Fu, M. L. Wong, and J. Licinio, "Genetic clustering of depressed patients and normal controls based on single-nucleotide variant proportion," *Journal of Affective Disorders*, vol. 227, pp. 450-454, 2018.

[136] S. Karthikeyan, and T. Christopher, "A hybrid clustering approach using artificial bee colony (ABC) and particle swarm optimization," *International Journal of Computer Applications*, vol. 100, no. 15, pp. 1-6, 2014.

[137] V. Kumar, J. K. Chhabra, and D. Kumar, "Automatic data clustering using parameter adaptive harmony search algorithm and its application to image segmentation," *Journal of Intelligent Systems*, vol. 25, no. 4, pp. 595-610, 2016.

[138] O. Botero, and H. Chaouchi, "RFID network topology design based on Genetic Algorithms," in *IEEE International Conference on RFID-Technologies and Applications*, pp. 300-305, IEEE, Sitges, Spain, 2011.

[139] O. Roeva, S. Fidanova, and M. Paprzycki, “Population size influence on the genetic and ant algorithms performance in case of cultivation process modeling,” in *Recent Advances in Computational Optimization*, Switzerland: Springer, 2015, pp. 107-120.

[140] N. Srinivas, and K. Deb, “Muiltiobjective optimization using nondominated sorting in genetic algorithms,” *Evolutionary Computation*, vol. 2, no. 3, pp. 221-248, 1994.

[141] F. Gu, H. L. Liu, and K. C. Tan, “A multi-objective evolutionary algorithm using dynamic weight design method,” *International Journal of Innovative Computing, Information and Control*, vol. 8, no. 5(B), pp. 3677-3688, 2012.

[142] H. R. Hassanzadeh, and M. Rouhani, “A multi-objective gravitational search algorithm,” in *International Conference on Computational Intelligence, Communication Systems and Networks*, pp. 7-12, IEEE, Liverpool, UK, 2010.

[143] X. Yuan, Z. Chen, Y. Yuan, Y. Huang, and X. Zhang, “A strength pareto gravitational search algorithm for multi-objective optimization problems,” *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 29, no. 06, pp. 1-39, 2015.

[144] H. Nobahari, M. Nikusokhan, and P. Siarry, “Non-dominated sorting gravitational search algorithm,” in *Proceedings of the International Conference on Swarm Intelligence*, pp. 1-10, Chongqing, China, 2011.

[145] G. A. Croes, “A method for solving traveling-salesman problems,” *Operations Research*, vol. 6, no. 6, pp. 791-812, 1958.

[146] A. Yadav, “AEFA: Artificial electric field algorithm for global optimization,” *Swarm and Evolutionary Computation*, vol. 48, pp. 93–108, 2019.

[147] K. Deb, A. Pratap, S. Agarwal, and T. A. M. T. Meyarivan, "A fast and elitist multi-objective genetic algorithm: NSGA-II," *IEEE transactions on evolutionary computation*, vol. 6, no. 2, pp. 182-197, 2002.

[148] B. Desgraupes, "Clustering Indices," *University of Paris Ouest-Lab Modal'X,* France, pp. 1–34, 2013.

[149] D. A. Van Veldhuizen, and G. B. Lamont, "On measuring multi-objective evolutionary algorithm performance," in *Proceedings of the 2000 Congress on Evolutionary Computation*, vol. 1, pp. 204–211, La Jolla, CA, USA, 2000.

[150] M. Halkidi, M. Vazirgiannis, and Y. Batistakis, "Quality scheme assessment in the clustering process," in *European Conference on Principles of Data Mining and Knowledge* Discovery, pp. 265–276, Springer, Berlin, Heidelberg, 2000.

[151] Z. Huang, "Clustering large data sets with mixed numeric and categorical values", in *Proceedings of 1st Pacific-Asia Conference on Knowledge Discovery and Data Mining*, pp. 21-34, Singapore, 1997.

[152] J. Ji, T. Bai, C. Zhou, C. Ma, and Z. Wang, "An improved K-prototypes clustering algorithm for mixed numeric and categorical data," *Neurocomputing*, vol. 120, pp. 590–596, 2013.

[153] C. Li, and G. Biswas, "Unsupervised learning with mixed numeric and nominal data," *IEEE Transections on Knowledge and Data Engineering*, vol. 14, no. 4, pp. 673–690, 2002.

[154] J. Ji, W. Pang, Y. Zheng, Z. Wang, and Z. Ma, "An initialization method for clustering mixed numeric and categorical data based on the density and distance," *International. Journal of Pattern Recognition and Artificial Intelligence*, vol. 29, no. 7, pp. 1-16, 2015.

[155] M. Thakur, S. S. Meghwani, and H. Jalota, "A modified real coded genetic algorithm for constrained optimization," *Applied Mathematics and Computation*, vol. 235, pp. 292-317, 2014.

[156] Y. Liu, B. Niu, and Y. Luo, "Hybrid learning particle swarm optimizer with genetic disturbance," *Neurocomputing*, vol. 151, no. 3, pp. 1237-1247, 2015.

[157] E. Zitzler, M. Laumanns, and L. Thiele, "SPEA2: Improving the strength Pareto evolutionary algorithm," *TIK-report*, vol. 103, pp. 1-19, 2001.

[158] M. A. Elaziz, N. Nabil, A. A. Ewees, and S. Lu, "Automatic Data Clustering based on Hybrid Atom Search Optimization and Sine-Cosine Algorithm," in *IEEE Congress on Evolutionary Computation*, pp. 2315-2322, Wellington, New Zealand, 2019.

[159] X. Li, "A non-dominated sorting particle swarm optimizer for multi-objective optimization," in *Genetic and Evolutionary Computation Conference*, pp. 37-48, Springer, Berlin, Heidelberg, 2003.

[160] M. A. Guzmán, A. Delgado, and J. De Carvalho, "A novel multi-objective optimization algorithm based on bacterial chemotaxis," *Engineering Applications of Artificial Intelligence*, vol. 23, no. 3, pp. 292-301, 2010.

List of Publications:

1. H. Petwal and R. Rani, “An Efficient Clustering Algorithm for Mixed Dataset of Postoperative Surgical Records,” *International Journal of Computational Intelligence Systems*, vol. 13, pp. 757-770, 2020. **[SCIE Indexed, Impact Factor - 1.8]**

2. H. Petwal and R. Rani, “Prioritizing the Surgical Waiting List-Cosine Consistency Index: An Optimized Framework for Prioritizing Surgical Waiting List,” *Journal of Medical Imaging and Health Informatics*, vol. 10, pp. 2876-2892, 2020. **[SCIE Indexed, Impact Factor - 0.6]**

3. H. Petwal and R. Rani, “An Optimized Framework for Surgical Team Selection,” CMC-Computers, Materials & Continua**, [SCI Indexed, Impact Factor – 4.89]**

4. H. Petwal and R. Rani, “An Improved multi-objective Water *Cycle Algorithm to modify inconsistent matrix in Analytic Hierarchy Process,” in Proceedings of Springer International Conference on Machine Intelligence and Data Science Applications (MIDAS),* University of Petroleum & Energy Studies, Dehradun, India, 4–5 September, pp. 187-197, 2020. **[Springer Conference]**

5. H. Petwal and R. Rani, “An optimal Multi-Criteria Decision-Making Framework to select best Multispecialty Hospital for surgery,” in *Proceedings of IEEE* 6^{TH} *International Conference on Parallel, Distributed and Grid Computing (PDGC),* Jaypee University of Information Technology, Waknaghat, Solan, HP, India, 6–8 November, pp. 471-475, 2020. **[IEEE Conference]**