European Journal of Operational Research 255 (2016) 593–603

Contents lists available at ScienceDirect

European Journal of Operational Research

journal homepage: www.elsevier.com/locate/ejor

Innovative Applications of O.R.

Dynamic resource allocation to improve emergency department efficiency in real time

Ruth Luscombe, Erhan Kozan*

School of Mathematical Sciences, Queensland University of Technology, GPOBox 2434,2 George Street, Brisbane, QLD 4001 Australia

ARTICLE INFO

Article history:
Received 21 January 2015
Accepted 20 May 2016
Available online 29 May 2016

Keywords:
OR in health services
Emergency department
scheduling
meta-heuristics
real-time

ABSTRACT

A dynamic scheduling framework is proposed to provide real-time support for managing the scarce resources of the Emergency Department. The theory of parallel machine and flexible job shop environments are integrated to schedule patient-bed assignments and task-resource allocations. The solution method incorporates dispatch heuristics, disjunctive graph methods and meta-heuristic search in order to provide fast solutions that respond to unscheduled arrivals, competing priorities and heterogeneous patient care needs. The dynamic algorithm is compared against static solutions and is shown to achieve solutions within 5 percent of the best bound. The dynamic schedule updates are completed within 2 seconds of information updates. This level of decision support, when implemented within a patient management system, can reduce the clinicians' workload by managing prioritized task lists such that the clinical staff is more free to focus on delivering clinical care.

1. Introduction

The Emergency Department (ED) is an important health care service that delivers time-critical care to unscheduled patient arrivals. Demand for ED services is increasing at a rate faster than population growth (FitzGerald et al., 2012) and bed availability (Forero et al., 2010). This motivates our focus on the effectiveness and efficiency of beds and resources within the ED. The key challenges for the ED environment are to accommodate unscheduled arrivals, competing priorities and heterogeneous patient care needs. The aim of this paper is to introduce a resource allocation model that represents the ED environment and can be implemented within a real time electronic patient management system.

Patients arrive at the ED and are triaged. The triage assessment determines the priority classification and immediate needs for each patient. The triage priority classes each have a recommended maximum response time to assign an ED bed or treatment space and commence ED care. The most urgent triage class has a recommended response time of zero and will bypass the triage stage for the immediate commencement of ED care. For all other triage priority classes, if the patient cannot be assigned to a treatment space then they wait until an appropriate treatment space becomes available. ED care starts with a medical assessment that develops into a treatment plan comprising of several treatment tasks. Each treatment task requires a staff or equipment resource for processing. The tasks include investigations such as imaging and pathology, treatments, and a disposition processes relating to the discharge pan for that patient. After ED care, the patient is discharged to their usual place of residence or admitted to a hospital unit for further care. Fig. 1 illustrates a basic model of patient flow within the ED.

Performance of the ED is measured primarily through time-based targets, such as response time (from arrival to bed assignment) and completion time (from arrival to discharge).

The ED is organized into different areas, each area contains several treatment spaces with an identical setup and is suited to different patient needs. Some areas have highly specialized equipment for monitoring and advanced life support, while other areas have a simpler setup used for low complexity care. For simplicity we will refer to the treatment spaces as bed resources, although some treatment spaces are fitted with beds and others (especially in the low complexity care areas) are fitted with chairs. Staff and equipment resources are organized into resource groups. All individuals within a resource group are assumed to be identical and skilled for a particular subset of treatment tasks. Doctors, nurses, and wardies are staff resources. Doctors and nurses are primarily responsible for clinical tasks relating to patient treatment and wardies are responsible for transporting non-ambulant patients to imaging appointments and for admission to inpatient wards. Pathology and X-ray are equipment resources used for diagnostic tasks.

The resource allocation model receives information from the triage assessment and medical assessment in the form of task

* Corresponding author. Tel.: +61 7 3138 1029; fax: +61 7 3138 2310.
E-mail address: e.kozan@qut.edu.au (E. Kozan).

http://dx.doi.org/10.1016/j.ejor.2016.05.039
0377-2217/

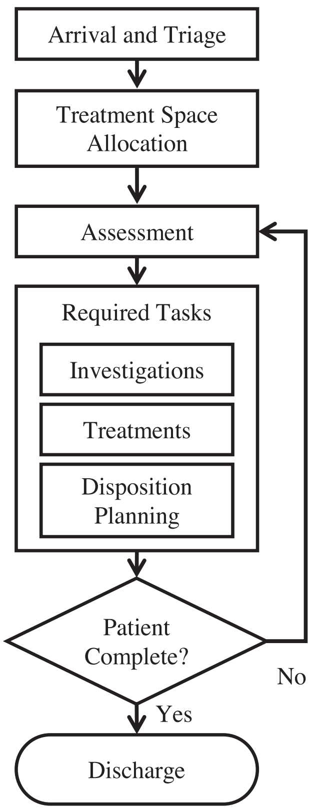

Fig. 1. ED patient flow.

requests. The bed assignments and treatment tasks are scheduled in such a way as to respect the patient priorities and maximize both resource efficiency and patient throughput. This model accepts the instructions of health care professionals in terms of patient priorities, bed requests, treatment tasks and gives as output an efficient resource allocation schedule.

The modeling for this research is based on a major regional hospital in Australia with over 45,000 emergency presentations per year. The ED has 30 treatment spaces with staff numbers between 15 and 30 depending on peak versus non-peak shift hours. Historical data from the ED information system is used to provide information regarding the arrival rates and the distribution of priority classes and other relevant statistics to describe the patient cohort.

The ED environment is modeled as a two-layer assignment and sequencing problem. The patient-bed assignment is modeled as a parallel machine scheduling environment with machine groups. The task-resource allocation is modeled as a flexible job shop. The two layers are interconnected since the bed assignment of each patient must span the duration of all treatment tasks.

The flexible job shop is defined by a set of resources and a set of jobs where each job consists of a set of precedence constrained tasks. Each task is processed by a single resource from an allowable group. In the ED setting patients are the jobs with precedence constrained treatment tasks. The resources are separated into distinct groups representing different kinds of staff and equipment resources e.g. doctors, nurses, X-ray, pathology, etc. The objective for the task-resource allocation is to minimize the total completion time for all patients.

The parallel machine problem with machine groups consists of a set of jobs and machines where each job must be assigned to an allowable machine for the duration of processing. In the ED each patient requires a bed from an appropriate subset of beds. The assignment of patients to beds occurs with the objective of minimizing weighted response time. The weights are determined from the priority class assigned to the patient at triage and the response time is defined as the time between the patient's arrival and their bed assignment time.

The resultant ED scheduling model is NP-hard since both the parallel machine environment and the flexible job shop environment are NP-hard (Moslehi & Mahnam, 2011; Pinedo, 2008). NP-hard problems are often unsuitable for exact algorithms due to the complexity and size of the solution space and so approximate algorithms are developed to find near-optimal solutions. We have developed a heuristic approach to generate solutions for the ED environment that exploits domain specific knowledge.

Since the arrivals of ED patients are not known in advance, the bed assignment is achieved using a priority dispatch rule. After the patient is assigned to a bed, the treatment tasks are scheduled using a disjunctive graph formulation for assignment and sequencing. The initial solution to the task-resource allocation is improved using a tabu search heuristic. When a new patient arrives, or when a bed becomes available for a new patient, these dynamic events trigger a new bed assignment. The treatment tasks for the new patient are incorporated into the disjunctive graph, which in turn leads to a new task-resource allocation and tabu search. This approach provides high quality solutions in a fast timeframe and allows for implementation in a dynamic environment.

This paper is organized as follows. Section 2 reviews the literature related to this problem. In Section 3 the model and notation is introduced followed by the description of the dynamic solution method in Section 4. The computational study is reported in Section 5 and the paper finishes with concluding remarks in Section 6.

2. Literature review

Quality of care in the ED is hard to measure in terms of patient outcomes as the confirmed diagnosis and the clinical follow-up are rarely made within the ED. It is further complicated because the three key stakeholder groups (patients, clinicians, and managers) have different perspectives on quality care (Cameron, Schull, & Cooke, 2011). Performance in the ED is often measured through time-based targets such as the response time (between a patients' arrival and the beginning of ED care) and completion time (from arrival to departure). In Australia the response time targets are defined using the Australian Triage Scale (ACEM, 2000) and the completion time targets are defined under the National Health Reform agenda (Australian Government, 2012).

There exist a large number of publications on modeling and analysis for the hospital in general and the emergency service in particular. Common quantitative approaches for health care tend to statistical analysis, simulation and mathematical modeling (Brailsford, Harper, Patel, & Pitt, 2009). Statistical analysis uses historical data to identify trends and patterns in the system. Studies of this sort examine system behavior and the effect of introducing new protocols. Simulation is often used as a planning tool as it permits experimentation without risk to patients. It can be limited by the detail and accuracy included in the simulation model but is very accessible with many off-the-shelf products available. Simulation models can also be used in conjunction with mathematical models (Diefenbach & Kozan, 2011). Mathematical models include queuing theory (QT) and mathematical programming models (MP). QT is used to plan system capacity and service levels to meet known patterns of demand. MP methodologies are used for scheduling and allocation of scarce resources, for example appointment scheduling, staff rostering, treatment task scheduling.

Treatment task scheduling in real time is the focus for this paper. By this we mean to create an algorithm where patients are