Reducing Physician Burnout and Costs in Outpatient Healthcare Settings via Advanced Analytics

by

Célia Escribe

B.S., Ecole Polytechnique (2019)

Submitted to the Sloan School of Management
in partial fulfillment of the requirements for the degree of

Master of Science in Operations Research

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2021

Author . . . .
Sloan School of Management
May 7, 2021

Certified by . . . .
Retsef Levi
J. Spencer Standish (1945) Professor of Operations Management
Thesis Supervisor

Accepted by . . . .
Patrick Jaillet
Dugald C. Jackson Professor, Department of Electrical Engineering and Computer Science
Co-director, Operations Research Center

Reducing Physician Burnout and Costs in Outpatient Healthcare Settings via Advanced Analytics

by

Célia Escribe

Submitted to the Sloan School of Management
on May 7, 2021, in partial fulfillment of the
requirements for the degree of
Master of Science in Operations Research

Abstract

National studies show that US primary care physicians are at high risk of burnout. Burnout has severe consequences for physicians, patients, and the healthcare system itself, which makes it one of the top priorities to be addressed by healthcare leaders today. In parallel, electronic health record systems (EHRs) have become ubiquitous in most practices, changing profoundly the work of primary care physicians, while offering at the same time new opportunities to analyze in a quantitative way various aspects of physicians' work. This thesis offers actionable insights driven by data and analytics to address the previously mentioned key healthcare challenges.

Chapter 2 leverages advanced text analytics to identify work themes of primary care physicians related to inbox message management. A scalable methodology relying on a Latent Dirichlet Allocation model is developed to analyze physicians' inbox work themes in great level of detail. This methodology could be used to implement appropriate workflow redesign in order to ensure that physicians spend most of their time on issues for which they have significant added value.

Chapter 3 employs a novel approach to examine team dynamics and their impact on physicians' well-being. While many studies have tried to isolate factors related to physicians' burnout, most of those studies are not scalable as they rely on self-evaluated surveys. This chapter addresses this question by providing a new quantitative methodology to analyze care team dynamics and structure through the integration of EHR data with social network modeling. Machine learning models are then developed to predict different dimensions of physicians' well-being using predictors related to team care dynamics and structure and work composition.

Chapter 4 finally develops a new modeling framework for a real-time appointment scheduling problem called the minimum peak appointment scheduling (MPAS) problem. While previous studies have applied existing algorithms from online bin packing to solve this problem, this modeling framework leverages unique aspects of appointment scheduling to further optimize scheduling decisions and reduce resource requirements. This chapter describes the first competitive online algorithm to the MPAS problem called the harmonic rematching (HR) algorithm, and proves that the

HR algorithm has an asymptotic competitive ratio of 1.5.

Thesis Supervisor: Retsef Levi
Title: J. Spencer Standish (1945) Professor of Operations Management

Acknowledgments

I would like to thank many people who have shared my path during the last two years I spent at MIT.

First and foremost, I would like to thank my advisor, Retsef Levi, for his mentorship and guidance over the past two years. His positive attitude and his aim at positive impact are striking traits of his personality that kept me motivated week after week. His optimistic attitude towards work has been incredibly inspiring to me. In addition to being an impressive researcher, Retsef is a caring advisor. I have been impressed at his capacity to always place his advisees on top of his priority list, which is always very full. His dedication to his work and his students has been a great source of inspiration to me.

My thesis focuses in part on work done in collaboration with faculty and staff at the Massachusetts General Hospital to whom I owe a great deal. I am especially grateful to Dr. Stephanie Eisenstat and Dr. Walter O'Donnell for their unwavering support and patience while working with me. I truly admire their dedication to not only caring for patients, but also to improving the well-being of their fellow physicians. They helped me understand better the world of physicians and all its challenges, and shared their will to improve the situation.

The past two years would not have been so enjoyable without the wonderful friendships I've developed at MIT and beyond. Although the global pandemic situation changed completely the experience, my first six months at MIT were a great opportunity to build many friendship bonds with the ORC crew, that I took care to attend to and develop in the next months. I would like first to thank Cynthia who has been a truly great friend during this time at MIT. I will remember with great pleasure our squash games and our fun dinners in Boston. I would also like to thank the French crew, Moise, Léonard, Amine. Thank to them, the MIT homework experience turned out to be funnier than expected. I also had the great pleasure to see my cousin Bianca arriving at MIT last September. After 20 years on two sides of the ocean, we finally got to spend some time in the same university.

I would finally like to thank my family, who has been a constant presence and comfort ever since I was young, and with whom I had the great pleasure to enjoy more time than expected over the last months, because of the sanitary crisis. Lastly, I could not have completed my Master's without the unwavering support of Gaspard.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The United States faces a large projected physician workforce shortage over the next decade, with the US Department of Health and Human Services projecting a shortfall of 54,100 to 139,000 physicians by 2033, and a shortfall of 21,400 to 55,200 primary care physicians by 2033 [6]. A variety of factors contribute to this workforce shortage, including population growth, an aging US population, a rapid increase in the complexity of medical care, an inadequate supply of physicians in medical school training, and mismatch between the residency positions available and the type of specialists needed [94, 93]. Physician burnout is another essential factor that may influence the adequacy of the physician workforce. National studies indicate that 43.9% of US physicians are experiencing professional burnout [90, 92]. Among all physicians, primary care physicians (PCPs) are especially at high risk, with more than 50% of PCPs reporting at least one manifestation of burnout. Fewer physicians are choosing primary care and many are leaving it [16, 101]. This could bear great consequences on the american healthcare system.

In parallel, electronic health record systems (EHRs) have become ubiquitous in most primary care practices and have changed consequently the work of PCPs. US physicians now spend numerous hours daily interacting with the EHR [8]. This surge of EHR has led to more work that is not direct face time with patients, including inbox management, computerized physician entry orders, structured documentation requirements and redistribution of some tasks which used to be performed by other clinical staff than physicians [98, 74, 67]. As part of the national healthcare reform effort, the federal government dedicated substantial resources to initiatives such as incentivizing “meaningful use” of EHRs. This initiative, which has a budget exceeding $35 billion [103, 80], was designed to improve the quality of care and increase coordination throughout health systems. However, many studies suggested that the widespread implementation of EHRs in the US may have resulted in significant unintended negative consequences impacting physicians [5, 1, 102, 51].

This raises a number of research and practical questions: (1) how to leverage EHR data to understand and alleviate PCPs’ work burden; (2) how to use this refined understanding of PCPs’ work to improve PCPs’ well-being; and (3) how to control costs and reduce resource requirements within health systems.

1.1 Leveraging Electronic Health Record Data to Analyze Physician Work

The dissemination of EHRs in almost all practices changed radically the nature of primary care physicians’ work. Traditional PCPs’ tasks such as refilling patients’ medication prescriptions and ordering laboratory tests which used to be relatively straightforward now often involve interacting with confusing user interfaces and navigating through multiple different EHR screens [17]. It was also highlighted that the surge of EHRs led to a shift in how workload is shared within the primary care team, increasing the administrative burden on physicians [51]. Physicians now spend a lot of time and effort to personally perform many tasks that do not require clinician-level training and do not contribute to building rapport with patients [3, 97]. At the same time, EHRs also became the major tool of communication between the primary care physician and patients as well as members in the supporting clinical care team [69]. This implies that patients are able to more easily contact their primary care physician through electronic channels such as online patient portals [72, 20]. This patient portal is used by patients to describe new symptoms, expect an opinion without presenting for a physical visit, or request a prior authorization for a medication [69]. While non-visit care has the potential of improving access to care by reducing office visits [25] and improving patient outcomes [4], such non-visit care also has the unintended effect to lead to a substantial physician workload that is often invisible. This finally implies more challenging offline care coordination across team members. These observations highlight the need to better understand the new nature of physician work, as this surge of EHRs has dramatically changed both the content of PCPs' work and the team dynamics in the PCP's work environment.

At the same time, EHR log data offers new exciting perspectives to analyze quantitatively various aspects of physicians' work. Core measures of EHR use have already been developed, such as total EHR time or 'work outside of work' [1]. Some studies focused on quantifying the work allocation between clinical face time with patients and time dedicated to EHR activities [5]. However, no study to the best of our knowledge has leveraged EHR data to analyze in a refined manner the content of physicians' work. In particular, the wealth of data on electronic inbox messaging between team members and with patients remains yet to be explored.

This study leverages advanced text analytics to develop a new methodology that can be implemented at scale to identify work themes of inbox messages managed by physicians on the EHR system. A Latent Dirichlet Allocation model is trained on inbox messages texts from a one-year period in order to identify different work themes managed by PCPs. This methodology then allows to describe each message as a mixture of different work themes, and therefore to calculate the relative share in workload of each work themes across physicians and clinics. The analysis identifies 30 different work themes. This large number and range of work themes which are managed by PCPs highlight the complexity of their work. This study also finds that on average, only 50.8% (range across physicians: 34.5 – 61.9) of the messages relate to medical tasks (e.g., diagnosis, care management) and 34.1% (range across physicians: 23.0 – 48.9) relate to administrative tasks (e.g., paperwork, scheduling). Therefore, a concerning fraction of physicians' time is spent on managing messages related to administrative matters that could be handled by other members of the clinic's team. This suggests that appropriate workflow redesign could better engage the care team to reduce the physician workload and ensure physicians spend most of their time on messages for which they have the maximum added value. This methodology could also be used to inform better understanding of the EHR inbox drivers of PCPs' burnout and work satisfaction and to optimize the care team work accordingly. Finally, variability at the practice level suggests a potential role for practice level redesign to improve the nature of primary care physicians' work.

1.2 Predicting Physician Well-Being

Burnout is a pervasive problem among physicians. The burnout syndrome is characterized by losing enthusiasm for work (emotional exhaustion), treating people as if they were objects (depersonalization), and having a sense that work is no longer meaningful (low personal accomplishment) [90]. The consequences of physician burnout are significant, and threaten the U.S. health care system, including patient safety, quality of care, and health care costs. Burnout appears to alter both the physician patient relationship and the quality of care physicians provide [45]. In particular, physicians' degree of burnout and professional satisfaction are related to physician empathy and compassion [89], prescribing habits [70], the risk of malpractice suits [10, 57], patient satisfaction [83], and whether or not patients adhere with physicians' medical recommendations [32]. Health care systems also suffer from physician burnout, as burned out physicians also have lower productivity in terms of number of hours worked and patients treated [17, 88, 31], and the costs of replacing a physician are estimated to be 2 to 3 times the physician's annual salary [19, 7].

While the problem of physician burnout has now been widely recognized, there is less information on how to address this problem. Some studies have tried to identify factors that contribute to physician burnout [94, 106, 89]. Teamwork has been in particular identified has a major driver of burnout. However, most of such existing studies rely on self-evaluated surveys where physicians report their satisfaction with different aspects of their work, and no study to the best of our knowledge has leveraged EHR data to analyze quantitatively teamwork. New scalable methodologies are therefore required to describe quantitatively the dynamics and structure of physicians' work ecosystems, and to relate them to physician well-being. Such methodologies should rely on the rich EHR datasets which are now easily available. Moreover, most quantitative studies trying to find associations between individual, work units or organizational characteristics and well-being have focused on a single component of well-being, burnout. However, physician well-being is a broader concept than burnout, with several additional dimensions worthy of assessment, including engagement, professional fulfillment, fatigue, emotional health, stress, and various dimensions of quality of life. There is an urging need to explore associations between other metrics than burnout describing physician well-being and various aspects of physicians' work ecosystem.

This study examines the topic of primary care physicians' well-being through the lens of teamwork, by exploring in a new quantitative way team relationships and their interplay with physician well-being. More specifically, this study leverages EHR data to provide a scalable methodology to analyze team structure and team dynamics in the primary care physicians' work environment. It utilizes social network research modeling to create features related to team interrelationships. This newly proposed methodology contrasts prior efforts in this space which have relied mostly on self-evaluated surveys where physicians report their satisfaction with different aspects of their work. This methodology offers by itself exciting perspectives as a new way to monitor quantitatively important aspects of physicians' work environment and experience. For example, it could be used to provide physicians with objective information on how their work experience compares with that of other physicians, and by healthcare leaders to compare different primary care practices' structure, dynamics and workflows. This methodology to describe teamwork is then combined with the methodology previously introduced in Section 1.1 to analyze the work content of EHR inbox messages. Regularized logistic regression models are then developed to predict various indicators traditionally used to measure physicians' well-being, and to identify underlying factors of physicians' well-being related to team care. The mean area under the receiver operating characteristic curve of the models over 1000 random partitions of the data into 80% training set and 20% held-out set for the selected dependent variables exhaustion, vigor, professional fulfillment and perceived appreciation is equal, respectively, to 0.701 [standard deviation (SD) 0.085], 0.692 [SD 0.085], 0.667 [SD 0.082] and 0.685 [SD 0.086]. The results show that factors related to the structure, dynamics and workflows of the supporting team of the physician are selected as important predictors of the different components of well-being. Moreover, it is found that the selected predictors vary across the different dependent variables. This highlights the complexity and multidimensionality of the concept of well-being. Examples of predictors associated with decreased well-being are the centrality of the physician in the support team in terms of flow of communication, the turnover within the support team, and the practice of forwarding messages without adequate information by nurses to physicians. Examples of predictors associated with increased well-being are a higher number of medical assistants in the support team graph and the share of messages related to ambiguous diagnosis. Therefore, this study allows to understand in a new refined manner physicians' team behavior and its interaction with physician well-being. Finally, this approach to team care could be used to improve systems and workflows in order to mitigate professional distress, as it paves the way for more in-depth exploration of operational system solutions which could be implemented.

1.3 Reducing Costs and Resource Requirements

Healthcare costs in the United States have continually risen over many decades. National healthcare spending has grown from $74.6 billion in 1970 to $1.4 trillion in 2000 to $3.5 trillion in 2017 [21]. There is therefore significant interest from health systems to control costs and reduce resource requirements.

This study addresses the challenge of reducing healthcare resource requirements by developing a new modeling framework for the ubiquitous problem of real-time appointment scheduling, called the minimum peak appointment scheduling (MPAS) problem. Whereas in most practical settings the appointment schedule is constructed by human schedulers who often are not able to systematically optimize the schedule, this study attempts to develop an optimization framework that leads to new algorithms with optimal theoretical performance guarantees and good practical performance. It is shown that the offline version of the MPAS problem is identical to the offline version of the well-known bin packing problem. However, in the online variant, there is additional flexibility that can be used to further optimize scheduling decisions and reduce resource requirements in health systems. This study describes the first competitive online algorithm to the MPAS problem called the harmonic rematching (HR) algorithm. The analysis shows that the HR algorithm has an asymptotic competitive ratio of 1.5. Considering that the current best lower bound on randomized online algorithms is 1.536, this highlights the fact that the HR algorithm performs better than any classical bin packing online algorithm in this setting. The strong empirical performance of the algorithm suggests several important insights that can inform scheduling approaches in practice.

1.4 Thesis

This thesis addresses the critical healthcare challenges described above, and provides actionable insights driven by data and analytics. Chapter 2 studies how primary care teams and patients e-communicate by identifying work themes via advanced text analytics. Chapter 3 provides a novel quantitative methodology integrating EHR data with social ntework modeling to explore the association of care team dynamics and work composition with primary care physician well-being. Chapter 4 finally describes a new modeling framework for the problem of real-time appointment scheduling, and introduces an online algorithm with theoretical guarantees for this problem.

Chapter 2

Identifying Work Themes in Electronic Health Record Inbox via Advanced Text Analytics

2.1 Introduction

Electronic health records have transformed the daily work of primary care physicians and are now ubiquitous in most primary care practices. In fact, it is known that PCPs spend more than half of their work on the EHR [5, 110, 102] with a significant fraction dedicated to clerical and administrative tasks [5, 110, 35]. Moreover, there are growing concerns that this phenomenon plays a major role in PCPs' burnout and attrition[1, 91, 8, 41]. Understanding the relationships between physicians' well-being and work in the EHR is critical if burnout is to be addressed more effectively. Partially motivated by this, there have been time motion studies and studies that leverage EHR log data to better understand various aspects of EHR work [42, 98]. Core measures of EHR use have been developed, such as total EHR time or 'work outside of work' (i.e., after formal hours) [99]. Other studies focus on quantifying the relative PCP work allocation between direct clinical face time with patients and time dedicated to EHR activities [98]. The desimination of EHR sytems has had opposing impact on work burden, reducing the burden on administrative support staff, and at the same time increasing PCPs' work burdens [51].

Inbox message management is an important component of EHR work burden, as it accounts for approximately a quarter of PCPs' workload [5]. EHR inbox has evolved over the last several years and became the major mode of communication between the PCP and patients as well as members in the clinical team. There have been only a few studies that have focused on this specific work component, most of which were done before the EHR inbox became a central mode of communication [76, 75, 68]. Others studies focused on quantifying number of messages received per relatively generic category (e.g., system-generated notifications, test results, communication with patient) [102, 74]. The goal of this study is to understand inbox work as a major communication tool of the PCP with patients and other care team members. In this context, it is critical to not only quantify the overall level of workload related to inbox message management, but also to understand in a more refined manner what PCPs are working on. This will be critical to identifying opportunities for workflow redesign that will hopefully help alleviating PCP's work burden.

To accomplish this goal, this study will focus on understanding related work themes (topics) in the messages managed by PCPs. Specifically, it is important to study what fraction of the EHR inbox messages directly relate to clinical diagnosis and management of patient health as opposed to administrative tasks (e.g., scheduling, paperwork) that can potentially be performed by other team members. Even with respect to clinical tasks, there could be opportunities to appropriately delegate them from the PCP to other team members. Additionally, it is insightful to understand how work composition varies across PCPs and clinics to identify modifiable drivers that can increase PCPs efficiency and work satisfaction.

One challenge to answer these questions is that it often requires massive manual review of messages or qualitative reporting by PCPs. This paper leverages advanced text analytics to develop a new methodology that enables analysis at scale of the work themes related to inbox messages managed by PCPs on the EHR. The results indicate that, in general, a substantial fraction of PCPs inbox management concern with messages related to administrative tasks. Moreover, there is significant variability among PCPs that seems to stem, at least partially, from practice level drivers (e.g., lack of protocol for handling electronic messaging) as well as individual work style.

2.2 Methods

2.2.1 Study Setting and Population

This work included a longitudinal analysis of all the EHR messages managed by 184 PCPs who work in 18 primary care practices at a large academic medical center in Boston, Massachusetts, from March 1, 2018 through March 1, 2019. Table 2.1 displays the characteristics of this cohort of PCPs. Note that the academic medical center has been using the Epic EHR system since 2016.

Table 2.1: Cohort Characteristics

Category	Physicians, No. (%)
Gender	
Female	118 (63)
Male	67 (36)
Years of practice	
10 – 19	37 (20)
20 – 29	52 (28)
30 – 39	36 (19)
≥ 40	14 (8)
unknown	17 (9)
Clinic Type	
Community Health Center	47 (25)
Boston Downtown	138 (75)
Clinical Full-Time Equivalent (FTE)	
0.9 – 1.0	20 (10)
0.75 – 0.89	21 (11)
0.50 – 0.74	53 (29)
0.25 – 0.50	70 (38)
< 0.25	22 (12)

2.2.2 Data Sources

The analysis uses a self-constructed dataset that integrates two main sources: (1) EHR inbox message data; and (2) PCP patient panel data.

In the academic medical center, the EPIC EHR system is de facto the primary messaging system for communication within the practice team and with patients. Inbox messages include messages from patients and other care team members, as well as messages generated automatically by the EHR system. Only the first category contains text information that is not protocol-driven. More specifically, those inbox messages include communications with the patient and between care team members initiated by a patient's message, a patient's phone call, transferring an electronical note related to a patient, or as part of the coordination of an action required for a given patient. The forthcoming analysis focuses on the latter category of patient-initiated inbox messages. Such patient-initiated communications corresponds to 36% of inbox messages received by the 184 PCPs included in the study. The remaining EHR protocol-driven messages can be divided into system-generated notifications, lab results notifications and prescription-related messages (authorization and signing), which represent, respectively, 26%, 9% and 29% of received inbox messages. This is summarized in Figure 2-1. The second main data source used in the analysis is an internal registry of patient assignments to the panel of individual PCPs. This defines the specific patient population for whom each individual PCP cares for.

Using the PCP patient panel data, each message from the inbox message data regarding a given patient is associated with a PCPs. Only inbox messages concerning patient-initialized communication are selected in the final dataset. This final dataset captures a total of 1,279,712 messages.

2.2.3 Analysis

Latent Dirichlet Allocation Model

A Latent Dirichlet Allocation (LDA) model was trained on the text messages of the self-constructed dataset [15]. The LDA model is a generative probabilistic model for

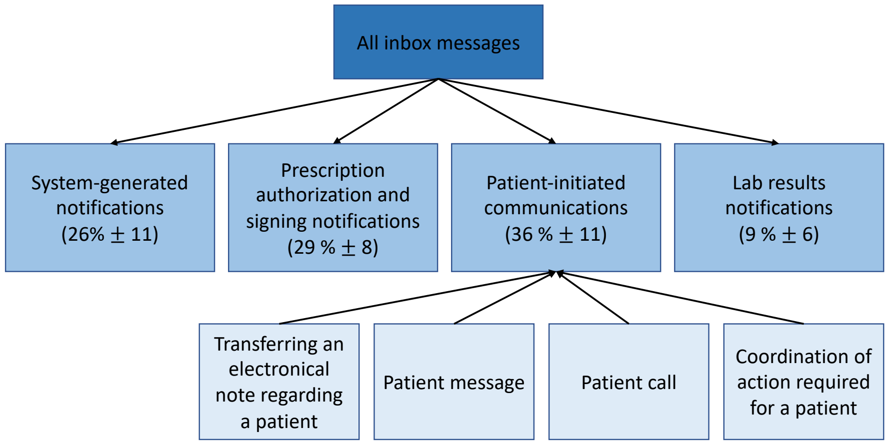

Figure 2-1: Inbox Message Allocation. The reported numbers correspond to the average number of received inbox messages allocated to a given subcategory, with the standard deviation across all PCPs. The bottow row corresponds to the different actions which can generate a patient-initiated inbox message.

discovering the topics that occur in a corpus of text documents. It is a three-level hierarchical Bayesian model, where the number of desired topics is given as input to the model. After training the model on all text messages, each identified topic is characterized by the respective frequency of each word from the corpus vocabulary. Each message is represented as a weighted mixture of the identified topics. For example, if the LDA model was trained to obtain 3 topics, then a given message can be categorized as the following mixture of topics: 60% topic 1, 40% topic 2, 0% topic 3. Topic 1 is considered the main topic of the message, while topic 3 is not present in the message. Due to uncertainty in the modeling, there is always a small percentage of the mixture of topics of a given message which corresponds to a none-specific topic.

The number of topics given as input to the LDA model was selected manually, retrospectively, among a list of potential number of topics (5, 10, 15, 20, 25, 30, 35, 40). Defining the number of topics by expert assessment is a common practice [48]. Indeed, likelihood-based comparisons of models on held-out sets only rely on the probability of observations, while the internal representation of the models is ignored. It was demonstrated that traditional metrics (e.g., perplexity) do not capture whether topics are coherent or not, suggesting that practitioners developing topic models should focus on evaluations that depend on real-world task performance rather than optimizing likelihood-based measures [23]. For each specified number of topics, an LDA model was computed. Each topic from the model and its corresponding word distribution was then explored manually. Topic meaning was subjectively interpreted from the topic-word distribution. Based on this interpretation, a topic was deemed coherent when its theme was salient and distinguishable. The quality of a model was then estimated based on the proportion of coherent topics among all topics. Finally, the LDA model with the highest proportion of coherent topics was selected as the best model, and its corresponding number of topics was selected as the input number of topics.

Work Theme (Topic) Identification

After selecting the best LDA model, a manual review was performed, with the help of two physicians, on each one of the topics identified by the LDA model in order to label the work theme captured in the respective topic. The first step of the manual review involved interpretation of the estimated word frequency within a topic. A general theme was selected based on this review. The second step of the manual review was validation by two physicians of the mixture of topics for 100 randomly selected messages. The methodology used to evaluate the accuracy of the predictions made by the algorithm for a given message is presented in detail in Appendix A.1.

Message Analysis

In order to understand the output of the model and the role of the different topics, it is important to explore if a message is most often described by a single topic, or by a mixture of different topics. To that end, all messages were reviewed by selecting their main (highest percentage) topic and the corresponding percentage. A relatively high percentage for the main topic of a message (e.g., 60%) implies that the message is primarily focused on this topic. On the contrary, a low percentage for the highest-percentage topic of a message (e.g., 20%) implies that the message includes a mixture of equally important topics. The statistics describing the percentage of the main topic across all messages were then calculated. The analysis was also performed at the topic level. For each topic, messages where the main topic corresponds to the given topic were selected, and statistics describing the main percentage were derived similarly.

Theme (Topic) Taxonomy

The overall goal of this methodological step is to provide a hierarchical taxonomy of the LDA-selected topics and the respective work themes, so that it is possible to analyze the share of PCPs' work concerned with more general work categories. The higher level of the proposed taxonomy includes *Administrative* and *Medical* issues. This fundamentally distinguishes between medical-related messages where the physician's input is expected to be needed compared to administrative-related messages which could be potentially handled by other care team members. A second more granular level was proposed to divide the *Administrative* category and the *Medical* category, respectively, into sub-categories. After discussion with two physicians, the *Medical* category was divided into the following categories: *Identified Symptoms, Ambiguous Symptoms, Condition Management, Tests and Exams* and *Clinical Decision-Making Referral.* Similarly, the Administrative category was divided into the following categories: *Paperwork, Scheduling, Administrative Referral* and *Prescription.*

Each previously identified message topic was then assigned to one of the categories of the proposed taxonomy described above. This final step was also performed with the help of two physicians.

PCP-level analysis

All statistics calculated at the PCP level considered all the patient-initiated messages where the PCP was involved either as sender or receiver of the message.

The share of messages concerned with a given topic was calculated for each PCP, by summing over all messages the respective percentage that this specific topic represented (ranging from 0% to 100%) and then calculating the mean. Following this analysis, it was possible to characterize, for each PCP, what the EHR messages related workload was across the different work categories in the taxonomy.

Practice-level analysis

Practice-level analysis was finally performed, by aggregating the statistics regarding the share of work per topic among all doctors from a given practice.

2.3 Results

Latent Dirichlet Allocation Model

Using human assessment of topic model validity as explained in Section 2.2 above, the best number of topics selected was 30. For models with smaller number of topics, some topics were missing, and other topics seemed to lump what would preferably be better differentiated themes. For models with larger than 30 topics, more topics were assessed to be less meaningful in terms of discerning a unique theme.

Work Theme (Topic) Identification

The top ten most frequent words for a selection of 9 topics along with hand-assigned labels are displayed as word clouds in Figure 2-2. The list of the top ten most frequent words for all topics is displayed in Table A.1. Sample messages and their topic mixture are presented in Appendix A.3.

The accuracy obtained with manual validation was equal to 73%, which corresponds to the average across all messages of the cumulative weight of the topics identified correctly. This indicates that most of the topics identified by the algorithm are correct. Moreover, this manual review revealed that the errors often occur with respect to topics with relatively small percentages, since the main topics of the mixture are mostly accurately identified by the algorithm.

Message Analysis

The mean percentage of the main topic across all messages was equal to 42.8% (standard deviation (SD) 16.6), and the median was equal to 40%. This implies that while it is uncommon to have a single topic representing a message, the main topic often

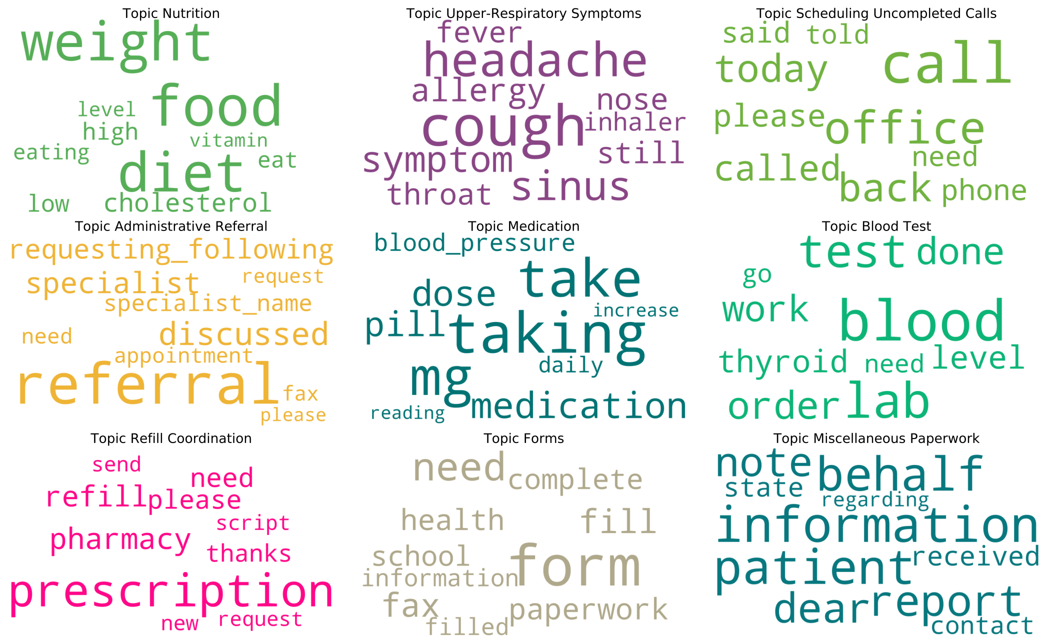

Figure 2-2: Word Clouds of the Top Ten Most Frequent Words for a Selection of Nine Topics.

has a significant weight in the message. The reported standard deviation implies that there is relatively high variability across the messages in terms of weight of the main topic. The same statistics at the topic level are represented in Table A.2. It should be noted that while some topics (e.g., *General Paperwork*) will often be the primary topic included in a message, others (e.g., *Musculoskeletal Symptoms*) will more often be discussed together other significant topics in a message.

Theme (Topic) Taxonomy

Each identified work theme (topic) was first assigned to one of the granular subcategories. For example, the following topics were assigned to the subcategory *Condition Management*: *Medication, Nutrition, Patient-Initiated Care, Chronic Cardiovascular and Diabetes Condition, Miscellaneous Chronic Condition, Joint Procedure* and *Surgical Procedure*. The topic assignment to one of the high level categories (i.e., *Administrative* and *Medical*) was then determined by the chosen subcategory. The complete resulting assignments of topics to categories are presented in Figure 2-3.

The same analysis regarding the percentage of the main topic was done for the

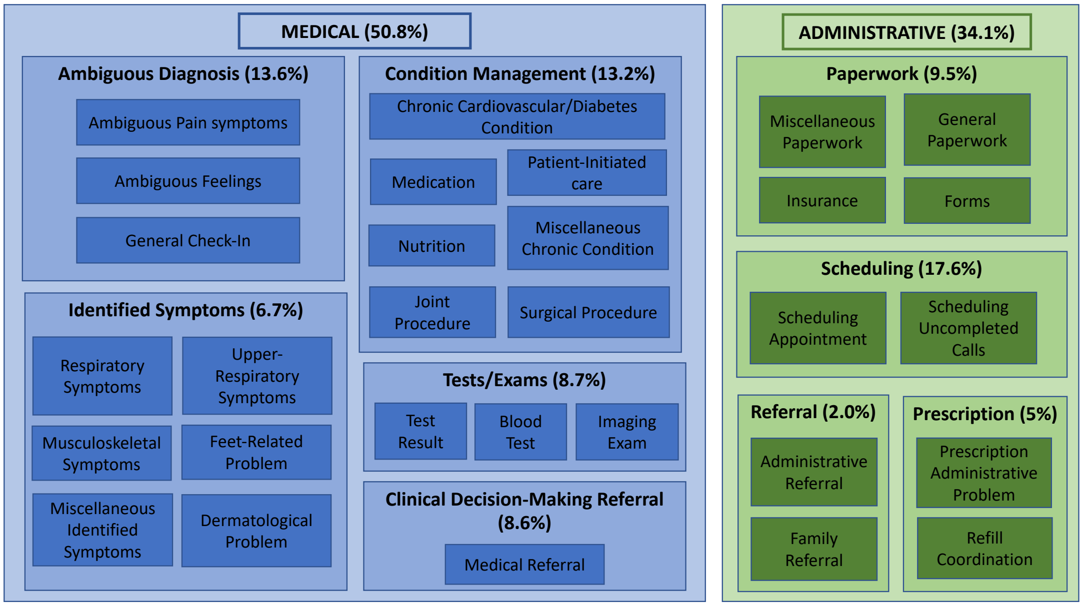

Figure 2-3: Topic Taxonomy for Patient-Initiated Inbox Messages Sent and Received by PCPs. Assignment of the 30 topics into the categories of the high-level taxonomy and the more granular taxonomy. Numbers displayed correspond to the average number of patient-initiated messages sent and received by PCPs concerned with each category across all PCPs by summing over all messages the respective percentage that this specific topic represented. *Medical* and *Administrative* categories amount to 84.9% of the messages, as the remaining 15.1% correspond to the average percentage of the unclassified topic in each message. The remaining protocol-driven inbox messages are not included in this analysis. The small value for the *Prescription* subcategory is explained by the fact that most prescription-related messages handled by the PCPs are protocol-driven ones (signing and authorization), and thus are not included in this analysis. The same holds for the *Referral* subcategory.

new classification. The mean percentage for the main topics when it belonged to the administrative work category was equal to 62% (SD 18). If the message belonged to the clinical work catergory the respective mean was equal to 70% (SD 18). This highlights that while messages can include multiple different topics, they are usually focused on one of the high level work category, specifically administrative or clinical.

PCP-level Analysis

On average, 34.1% (range: 23.0 – 48.9, SD 5.3) of the patient-initiated messages sent and received by physicians involved administrative issues, whereas 50.8% (range: 34.5 – 61.9, SD 5.3) of messages sent and received entailed medical issues. The other

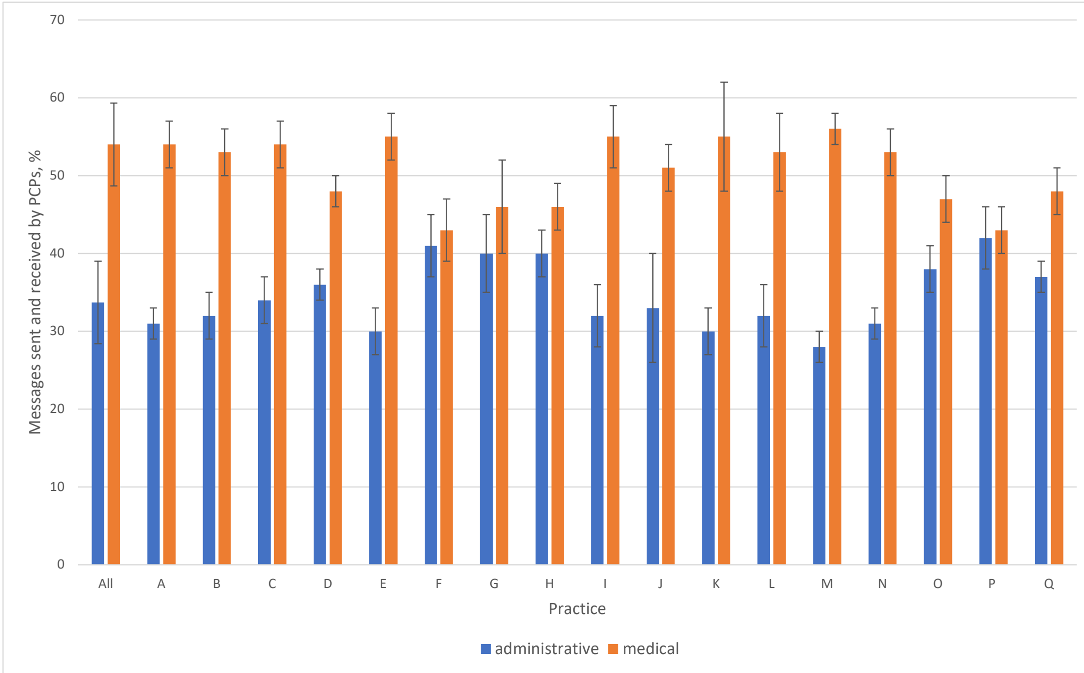

Figure 2-4: Share of inbox messages with standard deviation devoted to administrative and medical matters across all PCPs, and across PCPs from each one of the 18 considered practices.

15.0% (range 11.3 – 20.5, SD 1.7) were concerned with the none-specific topic. The large range indicates significant variability among PCPs. The average of the messages concerned with the subcategories is also presented in Figure 2-3.

Practice-level Analysis

Figure 2-4 reports the mean share of messages concerned with administrative and medical topics across different practices. The mean across all PCPs from a given practice differed greatly from one practice to another, ranging from 28% to 42% for administrative topic as an example.

2.4 Discussion

This paper illustrates that text analytics can provide a reliable and scalable methodology to identify work themes that constitute the PCPs' work related to EHR inbox patient-initiated messages management. This data-driven methodology allows to identify and characterize objective work themes instead of determining these categories rather subjectively. Large-scale data-driven analysis of the refined work themes is critical for the ability to re-design specific workflow processes to optimize the care team work, and particularly eliminate unnecessary workload on PCPs.

Since inbox messages is the primary mode of communication outside clinic visits, the analysis in this paper provides several indications that highlight the complexity of PCP work. The first indication is the vast number of topics which have to be addressed by the PCP. To add to this complexity, PCPs often have to manage both the clinical and administrative aspects related to the management of a patient. For example, handling a patient acute problem may require clinical diagnosis of the problem, as well as scheduling an appointment by the doctor.

This study found that a significant fraction of PCPs' time managing patient-initiated messages was spent handling administrative topics. Specifically, more than 34% of the meesages managed by the PCPs in the study were concerned with administrative topics. It is likely that some, if not the majority, of these administrative messages could be handled by other team members. For example, scheduling and paperwork issues mostly require the involvement of administrative assistants, and while signing on prescription may require the involvement of a PCP, administrative issues related to prescriptions (e.g., prescription sent to wrong pharmacy) could be addressed by nurses or medical assistants. While previous findings already highlighted physicians' administrative burden [5], no study has taken into account patient-initiated inbox messages when estimating the administrative burden. This study's findings are therefore all the more alarming, as they show that a significant part of physicians' work considered in the past as added value work is actually also driving administrative burden. Health care organizations may have to reconsider their conception of the role of EHR communication with patients and staff members in physicians' work burden.

This study also found a broad range of clinical work themes managed by PCPs. This raises the question of whether the PCP is always the best team member to handle all of them. For example, some of the messages related to the *Nutrition* topic could be delegated to nurses or medical assistants. Again, this suggests that health care organizations should take into account this great complexity of work themes currently managed by PCPs when working on team workflows redesign. Some studies suggested for example a shift from a physician-centric model to a shared-care model by expanding MAs' and nurses' roles to help meet patients' clinical needs with less use of the physician's time (e.g., with routine preventive counseling) [100].

The observed variability across practices in terms of the relative administrative share in the messages managed by the PCP suggests the hypothesis that practice level workflow processes, staffing schemes and overall supporting staff and infrastructure can have an impact on the PCP work composition.

The previous points suggest that the methodology presented in this paper allows to go further than merely quantifying the level of PCP work burden, and offer constructive directions to reduce this burden by redesigning inbox and other related practice workflows to allow PCPs focus their time on tasks that really require their clinical expertise. The complexity of the different work themes handled by the PCPs also suggest that a "one-size-fits-all" solution would probably fall short of success, and different workflows redesign should be considered depending on the specific inbox work theme. This study's methodology could be used to measure accurately a posteriori policy or workflow redesign impact on PCPs' work composition.

Finally, the improved understanding of PCP work themes and composition, as well as practice level factors could inform better understanding of the drivers of PCPs' burnout and work satisfaction, as well as mitigation strategies that are focused on changing the composition of their work and the manners by which it is conducted.

The main limitation of this study is that the methodology was used on PCPs within a single institution; thus, the generalizability of the obtained results to other hospitals is uncertain.

2.5 Conclusion

Handling EHR inbox messages is an important driver of work burden for physicians. This study found that a Latent Dirichlet Allocation model provides a reliable and scalable methodology to analyze the specific content of inbox messages. One of the main conclusion of this study was that content matters as much as quantity, as a consequent part of the time spent by PCPs interacting with other care team members and patients through the inbox was devoted to handling administrative matters. The great number and variety of topics managed by PCPs also highlighted the complexity of the nature of their work. Finally, variability across practices suggests that more effort should be devoted to understanding practice level drivers leading to increased administrative burden. This study therefore suggests that inbox workflows redesign could be implemented based on this topic taxonomy in order to alleviate PCPs' work burden by better engaging the whole team, leading to improved care efficiency.

Chapter 3

Integrating Social Network Modeling with Electronic Health Record Data to Predict Physician Wellbeing

3.1 Introduction

Physicians have very high rates of burnout (a syndrome characterized by exhaustion, cynicism and reduced effectiveness), with approximately half of them reporting at least one manifestation of burnout (43.9%-54.4%) [90, 92, 96]. Among all physicians, primary care physicians are especially at high risk of burnout [16]. Physician burnout has been shown to impact malpractice claims [24, 10], medical errors [71], turnover [108, 73], prevalence of substance use disorders [66, 78], and prevalence of depression and suicidal ideation [34, 37, 52]. The costs of replacing a physician if they decide to leave practice are estimated to be 2 to 3 times their annual salary [19, 7].

Many factors have been conjectured to contribute to physician burnout, including excessive workload, clerical burden, inefficiency in the practice environment, and the electronic health record [94, 95, 46, 36, 59]. While most existing studies rely on self-reported measures of satisfaction with work environment and EHR use [41], EHR log data has more recently been leveraged to quantify 'work after work', time spent on different tasks, and to find associations between such EHR use measures and self-reported burnout measures [5, 102, 1]. Most quantitative studies focus on burnout. However, physician well-being is a broader concept than physician burnout [106, 60]. There are several additional dimensions related to physician well-being worthy of assessment, including engagement, professional fulfillment, fatigue, and various dimensions of quality of life. [94]

It has been shown that interventions targeting structural factors within the work environment could reduce physician burnout [33], suggesting that the support team could be a potential driver of well-being [94, 107]. One study suggested that a shift from a physician-centric model to a shared-care model, with a higher level of clinical support staff per primary care physician and more standardized workflows, could result in improved professional satisfaction [100]. It has also been hypothesized that doing work not requiring clinician-level training may be associated with reduced PCPs well-being [100, 3]. These findings motivated analyzing the structure and dynamics of the support team and the themes of the work performed by PCPs when managing the EHR inbasket. To the best of our knowledge, this is the first study to provide a quantitative methodology to describe important aspects of the structure and dynamics of the PCPs' work environment and to relate them to physician well-being using machine learning.

This study develops a data-driven methodology based on social network modeling and EHR data to capture attributes related to structure and dynamics of PCPs' support team and work environment. Social network analysis [85] has been applied in different fields such as social media networks [50] or the study of working relationships [49] to investigate social structure through graphs. It characterizes networked structures consisting of nodes (individual actors) and the edges that connect pairs of nodes and typically correspond to the intensity of interaction between the respective two individuals. This approach is applied to the primary care practices to identify, for each PCP, the respective close *support team* with which the PCP works closely, including nurses, medical assistants (MAs), and front-desk staff members (FDs).

Additionnally, this study developed machine learning enabled analytical models to predict individual PCP metrics commonly used to measure well-being, including emotional exhaustion, vigor at work, professional fulfillment and perceived appreciation. This predictive approach illustrated that factors related to team structure and dynamics were important predictors for well-being.

3.2 Materials and Methods

3.2.1 Study Setting and Population

Physicians in an academic medical center (AMC) completed a well-being survey in May 2019 (overall response rate 92 %). Among the respondents, 251 PCPs answered the survey. EHR data was available for 188 PCPs from the AMC and their teams of nurses, MAs, FDs, and other support staff from 18 primary care practices from the AMC from March 1, 2018 through March 1, 2019. Survey data was available for 163 out of those 188 physicians (87% response rate). Demographic characteristics of the 188 PCPs are displayed in Table 3.1.

3.2.2 Data Sources

The analysis uses a self-constructed dataset that integrates 4 main sources: (1) EHR inbox message data; (2) results from the MGPO 2019 Physician Survey; (3) PCP patient panel data; and (4) data on work done after standard clinical hours.

The first data source was the EPIC EHR system which includes all communications between members of the PCP support team and with patients. For each message, the data includes the exact time of the message, the sender and the receiver of the message, the patient ID to which the message is related, and the text of the message itself.

The second data source consists of the individual results of the PCPs in the cohort from an internal well-being survey conducted by the AMC and distributed on May 20th, 2019. Included was the Maslach Burnout Inventory (MBI) [65, 9], the Professional Fulfillment Index (PFI), and the Utrecht Work Engagement Scale

Table 3.1: Cohort Characteristics

Characteristics		Respondents, No. (%)	Non Respondents, No. (%)
Gender	**Female**	105 (64)	14 (56)
	Male	58 (36)	11 (64)
Years of practice	$\leq$ **9**	19 (12)	5 (20)
	10-19	35 (21)	3 (12)
	20-29	54 (33)	3 (12)
	30-39	39 (24)	2 (8)
	$\geq$ **40** **unknown**	14 (9) 2 (1)	1 (4) 11 (44)
Clinical Full Time Equivalent (FTEs)	**0.90-1.00**	18 (11)	2 (8)
	0.75-0.89	18 (11)	3 (12)
	0.50-0.74	46 (29)	6 (24)
	0.25-0.5	66 (40)	6 (24)
	<0.25	15 (9)	8 (32)
Clinic type	**Community health centers**	46 (28)	4 (16)
	Boston downtown	117 (72)	21 (84)

(UWES) [86, 84].

The third data source used in the analysis is an internal registry of patient assignments to the panel of individual PCPs. This registry also contains the Health and Human Services-Hierarchical Condition Category (HHS-HCC) patient risk adjustment scores.

The last data source consists in the after-hour metric for each PCP, which corresponds to the percentage of each PCP's EHR activity that occurred during non-clinic hours, calculated with an existing methodology [53].

3.2.3 Teamwork Analysis Using Social Network Concepts

Social network concepts were used to create the support team graph and to analyze quantitatively the characteristics of each PCP's support team.

Support team graph The first step of the analysis determined the general practice team graph for each PCP. All inbasket communication messages involving the PCP's patient panel were considered. An undirected weighted graph was created based on those interactions. Each node in the graph corresponds to a member in the practice and there is a patient node that represents all patients belonging to the PCP's panel. An edge exists between two nodes in the graph if there is at least one communication documented between the two respective individuals. The weight of the edge corresponds to the number of such communications during the considered time period. The second step of the analysis created the support team graph, made of only team members that interact the most with the PCP through inbasket messages. All edges involving the PCP were ordered by decreasing weight, and the smallest set of edges in decreasing order containing at least 60% of the total number of communications of the PCP was selected. All staff members (and patient node) involved in the previously selected edges constituted the PCP's support team graph. The 60% threshold was selected by manual review with the help of two physicians. It was further validated by comparing the perceived close support team of 20 PCPs with the resulting support team graph. This methodology is represented graphically in Figure 3-1.

In order to capture the average teamwork characteristics over the one-year period, support team graphs were re-calculated for successive periods of 3 months with a rolling period of one month (e.g., March-May 2018, April-June 2018, etc... until January 2019-March 2019). Final features were obtained by calculating the average of a given feature over all the time periods.

Network-Based Teamwork Features The third step of the analysis was to calculate teamwork features based on the support team graph defined above with the aim to capture the structure and work dynamics of the support team, and the position of the PCP with respect to the support team. A complete list of features and their interpretation is presented in Table 3.2, and methodology to calculate some of them can be found in Appendix B.1. Selected features are presented here.

To capture the team composition in terms of roles (nurses, MAs and FDs) the

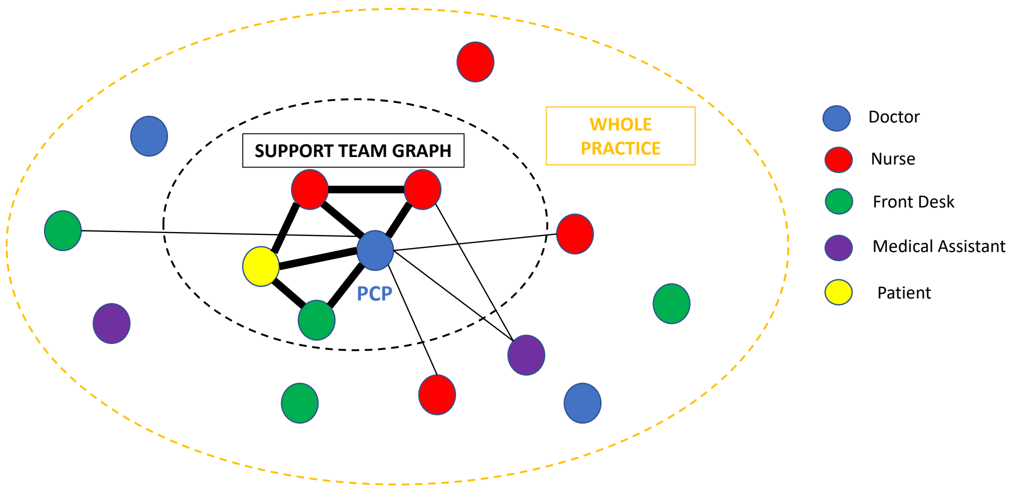

Figure 3-1: Schematic Representation of PCP Support Team Graph and Practice Graph. Nodes are staff members where the color represents the team member role. Edges represent inbasket communication between staff members and with patients. The weight of the edge is equal to the number of inbasket communications, and is represented by the width of the edge in this representation.

respective proportions of staff members with a specific role within the overall support team were calculated as features. The betweenness centrality was designed to capture the position of the PCP in the graph [77, 18]. This measures the centrality of the PCP in terms of flow of communication, namely to what extent can communication happen within the support team without involving the PCP. In order to capture the structure of communication within the support team, the entropy feature of the graph was calculated [30]. This feature was designed to describe how concentrated is the PCP's communication with the rest of the support team (e.g., concentration on a few heavy weight edges, or spread across many edges). Finally, turnover within the support team was estimated by comparing the PCP's support team from two adjacent time periods and calculating the number of staff members which left the team.

3.2.4 Inbox Work Content Analysis

Work Theme Inbox Analysis A descriptive methodology using advanced text analytics to identify work themes of PCPs related to inbasket management was developed in Chapter 2 and was used to derive features corresponding to the share of the

Table 3.2: Teamwork Features

Feature	Description
Number of nodes	How many staff members are included in the support team
Number of edges	How is communication organized among staff members in the support team
Weight	How much communication happens in the support team
Fraction of MAs	What is the proportion of staff members in the support team with a medical assistant role
Fraction of nurses	What is the proportion of staff members in the support team with a nurse role
Fraction of FDs	What is the proportion of staff members in the support team with an administrative assistant role
Ratio patient doctor	What is the proportion of communication with the patient handled by the PCP
PCP Betweenness centrality	How central is the PCP in terms of flow of communication in the support team
PCP Closeness centrality	How close is the PCP to all other staff members in the support team in terms of communication
PCP Entropy	How concentrated is the PCP's communication with the rest of the team
Turnover	Fraction of staff members who left the support team between two consecutive time periods

PCPs' inbasket work devoted to specific subcategories (e.g., scheduling, ambiguous diagnosis). Features related to such categories are presented in Table 3.3.

Inbox Writing Behavior The ways to communicate through inbox messages are also informative of teamwork dynamics. The PCP's writing behavior when writing to nurses and patients was described by the median length of messages written by an individual PCP to nurses and patients, respectively. The nurses' writing behavior when writing to the PCP was calculated in a similar manner. Nurses who forward many messages will have a lower median length of message to the doctor compared to more proactive nurses.

Table 3.3: Inbox Work Themes Features

Feature	Description
Scheduling	Share of messages sent and received by PCP concerned with scheduling matters (scheduling appointment, scheduling a call with a patient...)
Paperwork	Share of messages sent and received by PCP concerned with paperwork matters (insurance, health forms, authorizations...)
Prescription	Share of messages sent and received by PCP concerned with administrative prescription matters (refill coordination, problems with pharmacy...)
Administrative Referral	Share of messages sent and received by PCP concerned with administrative referrals
Identified Symptoms	Share of messages sent and received by PCP concerned with identified symptoms (acute respiratory problem, musculoskeletal symptom, dermatological problem...)
Ambiguous Diagnosis	Share of messages sent and received by PCP concerned with ambiguous diagnosis (ambiguous feelings, ambiguous pain symptoms...)
Condition Management	Share of messages sent and received by PCP concerned with condition management (chronic cardiovascular and diabetes condition, medication, nutrition...)
Clinical Decision-Making Referral	Share of messages sent and received by PCP concerned with clinical decision-making referral
Test and Exams	Share of messages sent and received by PCP concerned with tests and exams (test result, blood test, imaging exam)

3.2.5 Model Development

Dependent Variables Following existing approaches [1, 79], the PCPs were classified, based on their score for each of the wellbeing metrics, into two classes (0 or 1) based on a specified threshold. For example, physicians were classified as 1, i.e., low vigor, or 0, i.e., high vigor. Prior thresholds [65, 84] – when exist are based on overall population and do not necessarily capture the distribution of scores in a particular institution. To capture the specific distributions of scores on the studied cohort the median of the cohort was used as the threshold for the classification. Additionally, sensitivity analysis of the predictive models and performance with respect to the choice of the threshold was performed with thresholds ranging from 40%th quantile through the median.

Predictive Algorithm and Model Selection The implemented machine learning model was a Ridge regularized logistic regression algorithm [38]. Prior to training, highly correlated predictors with a Pearson coefficient > 0.75 were removed to reduce multicollinearity. Two parameters govern the regularization, specifically, the number of features and the ridge penalty parameter [39]. Regularization is used to decrease complexity of the model and to reduce overfitting. Few features and high penalty parameter lead to strong regularization and thus more robust predictive model. The classical machine learning approach of recursive feature elimination was implemented in order to select the right number of features, using the absolute value of the coefficient as the ranking criterion [39, 47, 111]. The number of features in the final model was selected by cross-validation. Penalty tuning was conducted via 10-fold cross-validation. The mean out-of-sample area under the receiver operating characteristic curve (AUC) along with the associated standard deviation (SD) were calculated via random partitioning of the data repeated 1000 times into 80% training set and 20% held-out set. The model inputs were finally restricted to the candidate predictors, and the final tuned model was fit over the entire training dataset to calculate the predictors' coefficients.

3.3 Results

3.3.1 Team Dynamics and Variability

Summary statistics for the different features introduced in the Methods section are presented in Table B.1. Very large standard deviations and ranges were obtained for the different teamwork features. For example, number of nodes, proportion of nurses in support team, turnover, betweenness centrality and entropy had respective average, standard deviations and interquantile range (IQ) across PCPs of 5.63 (SD 5.10, IQ 4.1 – 6.7), 0.44 (SD 0.45, IQ 0.36 – 0.51), 0.17 (SD 0.17, IQ 0 – 0.2), 0.67 (SD

0.73, IQ 0.38 – 0.86) and 1.18 (SD 1.15, IQ 0.87 – 1.48). Similarly, writing behavior was characterized by great variation, with average and standard deviation of number of words written by nurses to doctors equal to 18.1 (SD 11.4, IQ 10 - 25). Features describing allocation of inbasket work across subcategories presented smaller standard deviations, but ranges were still large. The relative share of messages handled by PCPs concerned with scheduling, ambiguous diagnosis and paperwork matters had respective overall range and interquantile ranges of 9%-27% (IQ 15%-19%), 7%-22% (IQ: 10%-16%) and 5%-29% (IQ: 8%-11%).

Statistics were compiled for the 25 PCPs who did not answer the well-being survey. No meaningful difference was observed with the rest of the cohort, as shown in Table B.2.

Great variability was finally observed at the practice level. Corresponding results are presented in Table B.3.

3.3.2 Model Performance for Predicting Well-Being

A model was built for each of the dependent variables. The results for the 4 dependent variables which yielded the most interesting insights are discussed in more detail here. Results for remaining dependent variables are presented in Table B.4 and Table B.5. The mean AUC of the models over the 80/20 random partitioning for the dependent variables emotional exhaustion, vigor, professional fulfillment and perceived appreciation was 0.701 [SD 0.085], 0.692 [SD 0.085], 0.667 [SD 0.082] and 0.685 [SD 0.086], respectively. Different features were identified as predictors for different dependent variables. Results are summarized in Table 3.4. Identified predictors of the different dependent variables included the turnover in the PCP's support team, the centrality of the PCP in the support team, the forwarding behavior of nurses, the fraction of medical assistants in the support team, the writing style of the PCP to the patient, the share of in-basket work pertaining to scheduling, ambiguous diagnosis and medical referral, FTE, mean risk score of the patients' panel, and after-hour.

Robustness of the results to the threshold was explored. The AUC score and the ranked coefficients of selected predictors were found to be robust to the choice of the

Table 3.4: Selected Predictor Variables for Regularized Logistic Regression Model.

Feature	**MBI exhaustion**	**UWES vigor**	**Stanford profes-sional fulfillment**	**Stanford perceived apprecia-tion**
FTE	0.91			
Graduation Year				-0.37
After Hour	0.50			
Risk score				0.95
Scheduling			0.43	
Ambiguous Diagnosis				-0.83
Medical Referral				0.48
Entropy	-0.75			
Fraction MA		-0.55		-0.36
Fraction nurse				
Weighted betweenness	0.53	0.47		
Turnover		0.34	0.10	
Length doctor patient	0.67	0.33		
Length nurse doctor		-0.47	-0.40	-0.35

threshold. Corresponding results are presented in Tables B.6, B.7 and B.8.

3.4 Discussion

The methodology developed in this study provides a novel approach to understanding the association between support team dynamics and structure and PCP well-being. The obtained features capture the variability in teamwork and how doctors interact with their work ecosystems. Variability in the number of nodes as well as in proportion of nurses, FDs and MAs indicates that PCPs use support teams with very different structure. It can be noted that there were in general few MAs in the support team graph, highlighting that most MAs are not highly involved in inbasket communication with the rest of the team. PCP's centrality and entropy were shown to vary greatly across PCPs. This highlights that there are very different support team dynamics and ways for PCPs to interact with their support team. Variation in allocation of inbasket work suggests that there is also variation in work sharing between the members of the support team. Variability across practices finally suggests practice level factors impacting teamwork dynamics.

Different factors were identified as predictors of dependent variables related to PCPs' well-being, highlighting that well-being is a complex multidimensional concept. PCP centrality was associated with high exhaustion and low vigor, while entropy was associated with low exhaustion. This suggests that workflows where communication is spread evenly across all staff members from the support team, and where staff members can coordinate care without directly involving the PCP are associated with improved burnout and engagement. This extends previous findings supporting a shift from a physician-centric model to a shared-care model [100], which had already suggested that empowering MAs and nurses by expanding their roles to help meet patients' clinical needs more effectively with less use of the physician's time could improve PCP satisfaction [97]. The fraction of MA's in the support team was found to be associated with high vigor and high perceived appreciation, while increase in nurses' message forwarding behavior to the PCP was associated with low vigor, low professional fulfillment, and low perceived appreciation; supporting prior findings. This highlights the positive association between PCPs' experience and expanding the support team to include more or more highly engaged MA's, as well as more proactive interactions from nurses when handling inbasket communication rather than forwarding for the PCP to manage the message.

Turnover was found to be associated with low vigor and low professional fulfillment. While it is known that burnout is correlated with turnover among PCPs [108], this finding indicates that turnover among support team staff also contributes to PCPs' engagement.

Allocation of inbasket work was selected as predictor of professional fulfillment and perceived appreciation, with the share of scheduling work associated with low professional fulfillment, and the share of ambiguous diagnosis associated with high perceived appreciation. Our findings suggest that scheduling tasks in particular are associated with decreased professional fulfillment, while solving complex diagnosis problems is positively associated with PCPs' perception of their work.

Finally, FTE and after hour work were associated with high exhaustion, while the average risk score of the patient panel was associated with low perceived appreciation. It can be noted that after-hour is only associated with exhaustion and no other well-being metrics. Interactions between dysfunctional teams and after-hour is an interesting future line of research, which can be undertaken using the methodology presented in this study.

The predictive results suggest that both the support team structure and dynamics and the content of PCPs' work are important predictors of PCP's well-being, and that PCP's well-being is not merely a function of PCP individual characteristics, but more a function of the overall system surrounding the PCP [94, 100]. For example, centrality of the PCP in the support team or the share of messages concerned with scheduling issues can be influenced by PCP's personal style, but more likely are affected primarily by team dynamics.

Deeper exploration of operational system solutions and potential impact on PCP well-being are necessary. Identifying predictors associated with components of PCP well-being can suggest specific directions for practice and workflow redesign. This methodology could also be used to provide PCPs with tools for self-calibration and objective data on how their work experience compares with that of other physicians.

There are limitations to this study. First, the model was trained on a relatively small sample of PCPs within a single academic medical institution. It would be interesting to observe the new insights derived from applying this methodology to a larger cohort, and to physicians from other specialties. Second, this study used cohort medians to define binary dependent variables instead of classical well-being thresholds because of the large bias between population distribution and classical thresholds. Therefore, this study's findings hold for a cohort of PCPs in a specific AMC and can inform operational solutions in this AMC, but they would not necessarily hold for PCPs in other medical centers. Third, the small sample size limited the range of predictive models which could be used, as larger sample sizes would allow to use more evolved models such as decision trees to capture nonlinear effects, or regression models to predict continuous dependent variables instead of binary variables. Lastly, as this analysis relies on inbasket communication, it underestimates physician work as it does not capture other additional face-to-face interactions.

3.5 Conclusion

Support team dynamics are critical to the work of physicians and their well-being. This study employs a novel approach to examine team dynamics and their impact on PCP well-being through a scalable and quantitative methodology using social network research concepts and available EHR data. Great variability was observed in terms of support team structure and dynamics across PCPs and practices. Factors related to the structure, dynamics and workflows of the PCP's support team were selected as important predictors of different dependent variables, highlighting that PCP's well-being is a function of the overall system surrounding the PCP, rather than features of the physician themselves. Future research should focus on operational system solutions relying on this new methodology.

Chapter 4

Competitive Algorithms for the Online Minimum Peak Appointment Scheduling

4.1 Introduction

This thesis chapter describes a fundamental online appointment scheduling model called the online *minimum peak appointment scheduling* (MPAS) problem. This model captures many practical settings of real time scheduling of appointments or jobs into a capacity limited resource. For example, consider an infusion center with a given number of infusion chairs. Patients' appointment requests for a given day arrive over time and each appointment request has to be scheduled upon arrival into a chair for a specified appointment duration. A natural objective would be to attempt to minimize the peak number of chairs occupied during the day, as the peak utilization typically dictates staffing levels, and could also affect how smoothly patients move through the center. Furthermore in many practical settings, once scheduled the appointment time cannot be easily changed.

Whereas in most practical settings the appointment schedule is constructed by human schedulers who often are not able to systematically optimize the schedule, this work attempts to develop an optimization framework that leads to new algorithms with optimal theoretical performance guarantees and good practical performance. To evaluate the performance of a newly developed online scheduling algorithm, the well-known concept of *competitive ratio* is employed. That is, the online algorithm performance is evaluated relative to an offline optimal algorithm that makes the optimal appointment scheduling decisions knowing all the appointments and their respective durations. The competitive ratio is defined to be the worst possible ratio between the respective objective values of the online algorithm and the offline optimal algorithm over all possible sequences of appointment requests and respective durations.

Interestingly, the offline version of the MPAS problem is identical to offline version of the well-known and well-studied *bin packing problem* [55, 29, 27]. In this problem, one has to pack items of varying length into bins of unit size with the goal of minimizing the number of bins used. Drawing on the appointment scheduling example above, the chairs are the equivalence of bins. It can be shown that minimizing the number of chairs needed to schedule ('pack') all the appointments is equivalent to minimizing the peak utilization. Moreover, any algorithm for the online bin packing problem can be applied to the online MPAS problem achieving the same competitive ratio. However, unlike the offline problems, the online bin packing problem is not identical to the online MPAS problem. The difference is that unlike the bin packing problem, where once an item is assigned to a bin, it cannot be moved to another bin, in the MPAS problem, the appointments are scheduled into time intervals, but there is flexibility regarding how to pack these time intervals into chairs. The interesting algorithmic question is whether this flexibility allows better performance. This work provides a positive answer to the question. In particular, the main result in the thesis chapter is a new online *harmonic rematching* (HR) algorithm for the MPAS problem that achieves an asymptotic competitive ratio of 1.5. This is strictly better than the best possible asymptotic competitive ratio on randomized online algorithms equal to 1.536 [22]. Moreover, the analysis indicates that the newly developed algorithm is 'almost' optimal by showing that no randomized online algorithm for the MPAS problem can obtain an absolute competitive ratio lower than 1.5. (The negative result is based on an example with finite number of jobs and does not entirely exclude that the lower bound might not hold as the number of jobs grows to infinity.)

To better articulate the comparison to the bin packing problem, this work will use interchangeably the terms appointment and jobs as well as chairs and bins. The newly developed algorithm for the MPAS problem directly leverages the flexibility mentioned above to pack the scheduled appointment duration intervals (jobs) into chairs (bins). It does that by scheduling what is called *one-sided bins*. Specifically, jobs that are scheduled consecutively starting from time 0 are called *left one-sided bins* and jobs scheduled consecutively and end at time 1 are called *right one-sided bins*. Moreover, the algorithm constructs one-sided bins of jobs that belong to the same job category. The job categories are defined in the spirit of the *harmonic algorithm* first developed in [61] for the online bin packing problem. The categories include *very large jobs* (duration over $\frac{2}{3}$), *large jobs* (duration between $\frac{1}{2}$ and $\frac{2}{3}$), *medium jobs* (duration between $\frac{1}{3}$ and $\frac{1}{2}$) and then multiple categories of small jobs that follow an harmonic pattern. For small and medium jobs, left and right one-sided bins of the same job categories are flexibly matched together into *complete sets of matched one-sided bins* and then further re-matched with large jobs to create *complete sets of scheduled bins*. The latter sets are used to create the schedule. The performance analysis of the algorithms relies on three important properties of the algorithm. First, the algorithm ensures that the complete sets of matched and scheduled bins have an average fullness of at least $\frac{2}{3}$. This is a desirable property since the average fullness provides a lower bound on the offline optimal solution. Second, the algorithm also ensures that the proportion of large jobs in the complete sets of scheduled bins is at least $\frac{2}{3}$. This property is desirable because one can have only one large job in each bin (chair) and therefore the number of large job provides yet another lower bound on the offline optimal value. The latter lower bound is particularly important when there is a high number of large jobs that cannot be matched with small or medium jobs. The third algorithmic property used in the analysis is that the respective duration distributions of the large and medium jobs placed on the right versus the left are not overly too imbalance. The latter property is achieved by implementing a randomization scheme to decide on what side to place arriving large and medium jobs. However, while the algorithm has a randomized element, the proven competitive ratio does not only hold on expectation, but occurs with probability 1 (as the number of jobs grows to infinity).

4.1.1 Contributions

The main contributions of this thesis chapter are as follows:

1. This study is the first to formally describe the MPAS problem. The objective of the model, minimizing the number of infusion chairs used, is aligned with real-world objectives of minimizing peak utilization. The new model captures fundamental practical scheduling settings.

2. Additionally, this study develops the first online algorithm for the MPAS problem, called harmonic rematching (HR) algorithm. Competitive analysis shows that the HR algorithm has an asymptotic competitive ratio of 1.5. Considering that the current best lower bound on any randomized online algorithm for the classical bin packing problem is 1.536 [22], this highlights the fact that MPAS and bin packing are fundamentally different and that the HR algorithm performs better than any bin packing online algorithm in this setting.

3. The strong empirical performance of the algorithm suggest several important insights that can inform scheduling approaches in practice. For example, current practices typically schedule appointments consecutively throughout the day, whereas the analysis in this study suggests that scheduling approaches that place appointments at the beginning and end of the day provide more flexibility to minimize the overall peak utilization. Another insight is that categorizing appointment into buckets of similar duration and consider that in devising the scheduling strategy might have significant benefits.

Table 4.1: Comparison of Bin Packing Setting and MPAS Setting.

Problem	Type of algorithm	Type of bound	Asymptotic ratio	Paper
Bin packing	Deterministic	lower bound	1.54278	[13]
Bin packing	Deterministic	upper bound	1.57829	[12]
Bin packing	Randomized	lower bound	1.536	[22]
MPAS	Randomized	upper bound	1.5	Ours

4.1.2 Literature Review

The importance of effective scheduling in healthcare contexts has been widely explored from both clinical and operational perspectives [62, 82, 87]. For example, scheduling for chemotherapy treatments was studied in [109] and [2]. These studies typically involve simulation models that help practitioners evaluate entire system processes under different scenarios and identify system bottlenecks. Some other studies also investigate methods for improving long-term treatment scheduling [104]. In contrast to these empirical studies, this study focuses on intraday appointment scheduling and incorporates theoretical analyses in addition to complementary empirical insights.

In the MPAS problem, the planner is tasked with scheduling appointment requests while trying to minimize peak utilization to avoid excessive strain on resources. As highlighted before, there are some essential similarities between the scheduling problem and classical *online bin packing*. In online bin packing, the decision-maker has access to an infinite supply of identically sized bins. The decision-maker receives items of various sizes one-by-one and must decide which bin to place each item into such that the sum of the sizes of all items in each bin cannot exceed the bin size. The goal of the decision-maker is then to place all of the received items into bins using as few bins as possible. The classical metric that is used to evaluate the performance of online bin packing algorithms is the *competitive ratio* [28, 27, 64, 26]. An online algorithm's competitive ratio measures the performance of the online algorithm by benchmarking it against an optimal offline algorithm, a solution that can be obtained if all the items are known in advance. Moreover, the competitive ratio is determined by the instance for which the ratio between the cost obtained by the online algorithm and the optimal offline algorithm is maximized. Competitive ratios are always lower bounded by 1, and lower competitive ratios are desirable. An online algorithm *Five-Thirds (FT)* was presented in [11] that achieves an absolute competitive ratio of $\frac{5}{3}$. This algorithm is actually optimal, since it is also proven in [11] that no online algorithm can have a lower absolute competitive ratio than $\frac{5}{3}$. A lot of interest has been focused over the years on the asymptotic behavior of online algorithms, which is captured through the asymptotic competitive ratio, when the number of items goes to infinity. There is extensive literature ranging back several decades examining the asymptotic competitive ratio of different online bin packing algorithms [29]. One of the earliest results dates back to the 1970's when it was shown that a well-known online bin packing algorithm called *Next Fit* had an asymptotic competitive ratio of 2 [54, 55]. This algorithm tries to pack the next item into the last bin that was used for packing, if such a bin exists and the item can be packed there, and otherwise it opens a new bin for the item. Since then, numerous other online algorithms have been developed and/or studied in the context of competitive ratios. For example, all bin packing algorithms belonging to a class referred to as *Almost Any Fit* have an asymptotic competitive ratio of 1.7 [55, 56]. A specific type of online bin packing algorithms are harmonic algorithms. In such algorithms, jobs are partitioned into categories by size. The simplest such categorization is based on harmonic numbers, leading to the *Harmonic algorithm* of [61]. A subset is defined as a category $(\frac{1}{j+1}, \frac{1}{j}]$. In these algorithms each subset is packed independently from other subsets using the algorithm *Next Fit*, meaning that the algorithm tries to pack a job in the last bin used for packing, and otherwise a new bin is opened for the job. For k growing to infinity, the resulting competitive ratio is approximately 1.69103 [61].

The current best upper bound was presented in [12], with the *Advanced Harmonic* algorithm which obtains an asymptotic competitive ratio of 1.57829. For the standard bin packing problem, asymptotic lower bounds have also been improved over the years [63, 105, 40, 14]. The current best known lower bound on the asymptotic competitive ratio is 1.54278 [13]. It was also shown in [22] that an asymptotic lower bound of 1.536 holds for any randomized online algorithm for bin packing. Those different results on the bin packing algorithms performance are summarized in Table 4.1.

The chapter is organized as follows. Section 4.2 first introduces and explains the MPAS problem, and derives a lower bound on any online algorithm for the MPAS problem. Section 4.3 then defines the newly proposed HR algorithm with the specific matching and rematching rules. The performance of this algorithm is analyzed in Section 4.4 through competitive analysis, and it is shown that the HR algorithm has an asymptotic competitive ratio of 1.5. In Section 4.5, experiments show that the HR algorithm achieves near-optimal performance in practice on generic problem instances.

4.2 Problem Definition

4.2.1 Model Development

This section formally defines the MPAS problem and describes the model notation. The problem focuses on scheduling appointments in a specific scheduling interval. Without loss of generality, it is assumed that the given scheduling interval has length $T = 1$.

There are N appointments denoted by $A_1, \dots, A_N$ with respective required scheduling times $l_1, \dots, l_N$ $(0 \leq l_i \leq 1)$, arriving one after the other. The planner schedules these appointments in real time upon arrival into a pool of resources (e.g., chairs). Specifically, let s_i be the decision to schedule the appointment A_i to start at a feasible time, with $0 \leq s_i$ and $s_i + \ell_i \leq 1$. For a given schedule $(\boldsymbol{A}, \boldsymbol{\ell}, \boldsymbol{s}) = \{(A_1, \ell_1, s_1), \dots, (A_N, \ell_N, s_N)\}$, the load at time t is equal to $u_t = \sum_{i=1}^{N} \mathbb{1}\{s_i \leq t < s_i + \ell_i\}$, which corresponds to the total number of appointments being conducted at time slot t. The peak load (maximum number of chairs) for a given schedule is $p((\boldsymbol{A}, \boldsymbol{\ell}, \boldsymbol{s})) = \max\limits_{0 \leq t \leq 1} u_t$. The planner's goal is to schedule the appointments so that the peak load of the resulting schedule is minimized.

The assumption is that at the time when s_i is determined, no information is provided on future appointments $i+1, \dots, N$. In light of the inherent online nature of the central planner's decision-making process, it is natural to consider online algorithms, and to measure their performance through the well-known concept of *competitive ratio*. For an online algorithm ALG and an input sequence of appointments I, let $ALG(I)$ be the maximum number of chairs used by the algorithm. The maximum number of chairs used by the optimal offline algorithm OPT is similarly defined as $OPT(I)$. The absolute competitive ratio is then defined as

$$\sup_{I \in \Sigma} \frac{ALG(I)}{OPT(I)}$$

where Σ represents the space of all sequences of appointments. Another common performance metric is the asymptotic competitive ratio which is defined as

$$\lim_{N \to \infty} \sup_{I : OPT(I) \geq N} \frac{ALG(I)}{OPT(I)}$$

The planner's goal is therefore to design an online algorithm with the smallest possible competitive ratio. The analysis developed in Section 4.4 will focus on the asymptotic competitive ratio, when N grows to infinity.

The following arguments show that any schedule $(\boldsymbol{A}, \boldsymbol{\ell}, \boldsymbol{s})$ can be executed such that each job is processed entirely by the same unit of resource (chair), and such that there exists a feasible appointment-to-chair assignment for the N appointments that uses exactly $p((\boldsymbol{A}, \boldsymbol{\ell}, \boldsymbol{s}))$ chairs.

Proposition 1. *For any feasible schedule $(\boldsymbol{A}, \boldsymbol{\ell}, \boldsymbol{s})$, there exists a feasible appointment-to-chair assignment for the N appointments that uses exactly $p((\boldsymbol{A}, \boldsymbol{\ell}, \boldsymbol{s}))$ chairs, and in which each appointment is processed by a single chair.*

Proof. Consider the interval graph corresponding to the given schedule [58]. This is the graph that arises when there is a node for each appointment interval $[s_i, s_i + \ell_i]$ and when there is an edge between two nodes if and only if the corresponding intervals overlap. Then finding a feasible appointment-to-chair assignment is nothing else but finding a coloring of the interval graph where the number of colors used is equal to the number of chairs used. The assignment using the smallest number of chairs is then equal to the chromatic number of the interval graph. On the other hand, consider a clique in the interval graph, which corresponds to a set of pairwise overlapping intervals. Then by the Helly property for interval graphs [44], there is one point where all of the intervals overlap. This implies that the peak load $p((\boldsymbol{A}, \boldsymbol{\ell}, \boldsymbol{s}))$ is equal to the size of the maximal clique in the interval graph. Now, it is known that an undirected graph is an interval graph if and only if G is a triangulated graph and its complement is a comparability graph [43]. This implies that interval graphs are perfect graphs [44]. Therefore, the chromatic number of the interval graph is equal to the size of the largest clique of the graph. This implies that there is an appointment-to-chair assignment that uses exactly $p((\boldsymbol{A}, \boldsymbol{\ell}, \boldsymbol{s}))$ chairs, and in which each appointment is processed by a single chair. □

Proposition 1 proves that the goal of the central planner to minimize the peak load is equivalent to minimizing the maximum number of chairs used to schedule all appointments, such that each appointment is processed by a single chair. Considering appointments as jobs, length of appointment as job size, and chairs as bins, it is then easy to see that the offline version of the problem studied in this work is equivalent to the well-known bin packing problem. This implies that any online bin packing algorithm will provide a feasible solution to this scheduling problem with the same competitive ratio. However, the added flexibility to delay the appointment-to-chair assignment procedure at the end of the scheduling procedure implies that classical online bin packing algorithms may not be optimal in this setting. Example 1 illustrates that the online versions of the two problems are likely to be different. Specifically, there is benefit from delaying the job-to-chair assignment.

Example 1. *Let $T = 5$ and suppose the planner must initially schedule four consecutive appointments with required scheduling times $l_1 = 3$, $l_2 = 1$, $l_3 = 2$, and $l_4 = 2$. These first four appointments could feasibly be scheduled with start times $s_1 = 0$, $s_2 = 4$, $s_3 = 0$, and $s_4 = 3$. If the chair assignments were performed at the time of appointment scheduling, it would naturally result in A_1 and A_2 being placed in one chair while A_3 and A_4 are placed in another chair. Suppose now that a fifth appointment request with required scheduling time $l_5 = 2$ is realized. Figure 4-1 clearly shows* *that this new appointment does not fit in either of the two chairs given the pre-existing chair assignment. A third chair would therefore need to be introduced. However, if the chair assignment had been delayed, then the chair assignments of A_2 and A_4 could have been swapped as shown in Figure 4-2 thereby enabling Chair 2 to accommodate A_5 and not necessitating the use of another chair, as shown in Figure 4-2.*

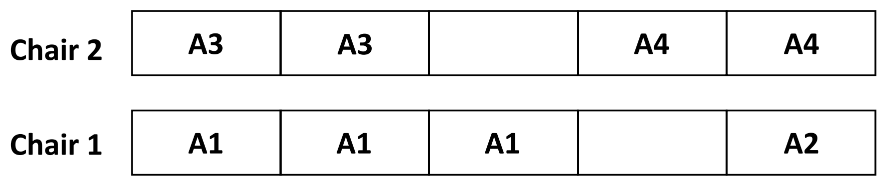

Figure 4-1: Real-time Chair Assignment Example.

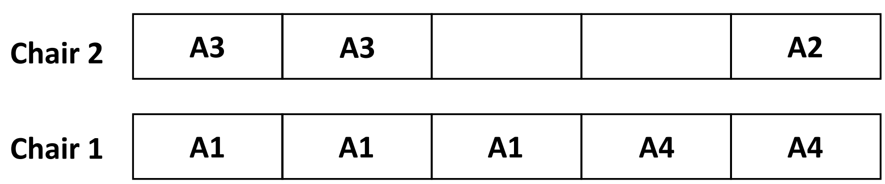

Figure 4-2: Delayed Chair Assignment Example

4.2.2 Lower Bound on the Absolute Competitive Ratio of any Online Algorithm

In this section, a lower bound on the absolute competitive ratio of any online algorithm for this scheduling problem is derived.

Proposition 2. *Any randomized or deterministic online algorithm for the MPAS problem has an absolute competitive ratio $CR \geq 1.5$.*

Proof. See Appendix C □

4.3 Harmonic Rematching Algorithm

This section introduces the HR algorithm which will subsequently be shown to have an asymptotic competitive ratio of $^3/_2$ for the considered scheduling problem.

4.3.1 Overview of the Algorithm

The general ideas driving the construction of the HR algorithm are presented in this section.

The first central idea underlying the proposed algorithm is to partition the jobs into categories with scheduling requirements within a specified range following the general form $(\frac{1}{j+1}, \frac{1}{j}]$ for some $j \in \mathbb{N}$. This follows in spirit the *Harmonic bin packing algorithm* [61]. The second idea concerns the way of assigning jobs to bins. Example 1 illustrates the potential flexibility allowed by the scheduling problem studied in this work compared to the classical bin packing problem. To leverage this flexibility, the algorithm described below will maintain what shall be subsequently called *left and right one-sided bins*. A *left one-sided bin* corresponds to consecutive scheduled jobs that start at time 0 and end at some time $0 < t < 1$. Similarly, a *right one-sided bin* corresponds to consecutive scheduled jobs that start at some time $0 < s < 1$ and end at time 1. Thus, *opening* a left one-sided bin corresponds to scheduling a job to start at time 0, and similarly *opening* a right one-sided bin corresponds to scheduling a job to end at time 1. The value of maintaining one-sided bins is the flexibility the algorithm has to dynamically match and rematch left and right one-sided bins together into a bin, as long as their total scheduled time does not exceed 1. The *fullness* of a bin is defined as the sum of the lengths of all the jobs in that bin. Since the lengths of all bins are normalized to 1, the fullness of a bin is equal to its utilization. Similarly, the *fullness* of a one-sided bin is defined as the sum of the lengths of all the jobs in that one-sided bin. The *average fullness of a set of bins* is defined as the average of the fullness of all the bins constituting this set.

The two ideas described above allow the algorithm to maintain appropriate balance between two metrics, fullness and rematching flexibility. In other words, the algorithm attempts to secure sufficient fullness of the matched one-sided bins, while ensuring at the same time sufficient flexibility to rematch appropriately defined one-sided bins together.

Finally, the algorithm employs some randomization when opening a new one-sided bin to decide whether it will be right or left, with the goal of maintaining certain properties related to the size distribution of one-sided bins on the right and left side and which turned out to be important for the competitive ratio analysis in Section 4.4.

4.3.2 Creation of One-Sided Bins

Job Category.

Following previous works [61, 12], jobs are partitioned into categories based on their sizes. A *very large job* is a job with length $\ell_i \in (\frac{2}{3}, 1]$. A *large job* is a job with $\ell_i \in (\frac{1}{2}, \frac{2}{3}]$. A *medium job* is a job with $\ell_i \in (\frac{1}{3}, \frac{1}{2}]$. A *small job* is a job with $\ell_i \in [0, \frac{1}{3}]$. Those small jobs are further partitioned into subcategories $I_j = (\frac{1}{j+1}, \frac{1}{j}]$ with $3 \leq j \leq 11$. The largest index category of small jobs is of the form $[0, \frac{1}{12}]$.

Types of One-Sided Bins.

One-sided bins are all category-specific in that each one contains jobs from only one category. For very large jobs, large jobs and medium jobs, one-sided bins are all made of a single job. Those one-sided bins will be called respectively *very large one-sided bins, large one-sided bins and medium one-sided bins*. Two one-sided bins are said to be *opposite* if they are placed on opposite sides. For small jobs, three types of one-sided bins can be distinguished. *Type 1 one-sided bins* are filled with the maximum number of jobs with total scheduled time t smaller than 1. For example, a Type 1 one-sided bin from category $(\frac{1}{6}; \frac{1}{5}]$ is made of 5 jobs, since the maximum number of jobs from this category that can be matched to a single bin is equal to 5. A Type 1 one-sided bin will also be called a *Type 1 bin*, since it will never be matched to another one-sided bin on the opposite side. There are also *Type 2 thin one-sided* *bins.* Type 2 thin one-sided bins are filled with the maximum number of jobs with total scheduled time t smaller than $\frac{1}{3}$. The last category of one-sided bins is called *Type 2 fat one-sided bins.* Type 2 fat one-sided bins are filled with the maximum number of jobs with total scheduled time t smaller than $\frac{1}{2}$. For example, a Type 2 thin one-sided bin from category $(\frac{1}{6}, \frac{1}{5}]$ is made of one job, and a Type 2 fat one-sided bin from category $(\frac{1}{6}, \frac{1}{5}]$ is made of 2 jobs. A precise definition of the different types of one-sided bins can be given for each category, as specified in Appendix D. For all categories of small jobs but the last one, the number of jobs in a Type 2 thin or fat one-sided bin can be specified precisely thanks to their scheduling requirements within a given range $(\frac{1}{j+1}, \frac{1}{j}]$.

Complete Set of Matched One-Sided Bins.

As jobs arrive, the algorithm creates one-sided bins. Central to the process is the notion of *complete sets of matched one-sided bins.* These sets are category-specific in that each one contains jobs from only one category. They are defined for small and medium jobs only. The intuition underlying the creation of complete sets of matched one-sided bins is the attempt to ensure fullness of existing matched one-sided bins, and at the same time to maintain flexibility.

First note that due to the way they are constructed, a Type 2 thin one-sided bin can always be matched to a Type 2 fat one-sided bin from the opposite side to form what shall be called a *Type 2 bin.* Moreover, two Type 2 bins are said to be *symmetrical* if their respective Type 2 thin one-sided bins are placed on opposite sides (see Figure 4-3). For small jobs ($l \leq \frac{1}{3}$), a complete set of matched one-sided bin is created by combining Type 1 bins and Type 2 bins, where Type 2 bins are obtained by temporarily matching Type 2 fat one-sided bins with Type 2 thin one-sided bins as described previously. Figure 4-4 shows an example of such a complete set of matched one-sided bins. This matching will potentially be modified during the scheduling process described in Section 4.3.3. A complete set of matched one-sided bin contains n_1 Type 1 bins and n_2 Type 2 bins where n_2 is even. Type 1 bins are used to increase the fullness of the complete set of matched one-sided bins, while Type 2 bins are used to protect against the arrival of large jobs.

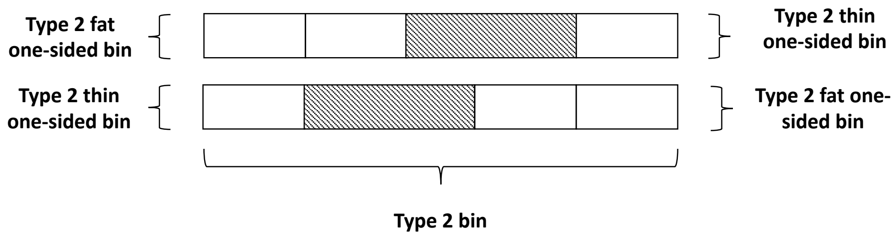

Figure 4-3: Two Symmetrical Type 2 Bins. Their Type 2 thin one-sided bins are placed on opposite sides. The grey parts represent unoccupied space in the bins.

A complete set of matched one-sided bins with n_1 Type 1 bins and n_2 Type 2 bins has the following properties:

- There are n_2 Type 2 thin one-sided bins, and n_2 Type 2 fat one-sided bins following the definition of a Type 2 bin.
- There is the same number of Type 2 thin one-sided bins on each side (left and right), and of Type 2 fat one-sided bins on each side (left and right).

The exact composition of a complete set of matched one-sided bins is specified for each job category in Appendix D. The proportion of Type 1 and Type 2 bins in each job category is defined so that a complete set of matched one-sided bins is at least $\frac{2}{3}$ full for each job category as will be proven subsequently in Lemma 1.

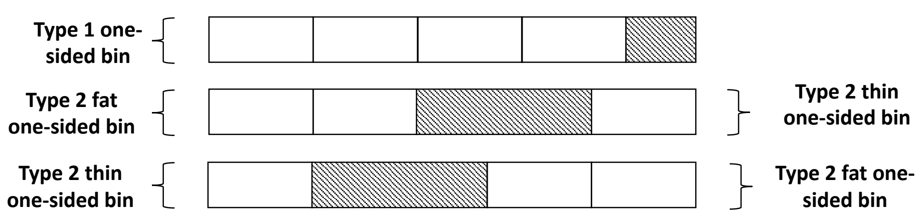

Figure 4-4: A complete set of one-sided bins from category $(\frac{1}{5}, \frac{1}{4}]$ with $n_1 = 1$ and $n_2 = 2$. There is one Type 1 one-sided bin containing 4 jobs, and two symmetrical Type 2 bins. A Type 2 bin is obtained from matching a Type 2 thin one-sided bin made of one job with a Type 2 fat one-sided bin made of two jobs. The grey parts represent unoccupied space in the bins.

For medium jobs, a complete set of matched one-sided bins is made only of two opposite medium one-sided bins matched together, and therefore only contains one type of bin.

A one-sided bin is said to be *incomplete* if it has not been filled by the number of jobs defined by its job category and type as specified in Appendix D. It is said to be *complete* otherwise. A set of matched one-sided bins is said to be *incomplete* as long as it does not meet all the criteria to form a complete set of matched one-sided bins from its job category as specified in Appendix D. It is said to be *complete* otherwise.

Arrival of a New Job.

As jobs arrive, the algorithm attempts to form complete sets of matched one-sided bins as described in Section 4.3.2. Each time that another respective set of matched one-sided bins is completed, the algorithm forms a new set. When a new job arrives, the following rules define what one-sided bin the new job will be part of:

- If it is **a very large or large job**, a new one-sided bin is created with this job. This job could be used to create either a left or right one-sided bin. If there is a smaller number of one-sided bins from this category on one side than on the other, the new one-sided bin is created on this side. Otherwise, the right or left side are chosen with equal probability of 0.5. This implies that if one large or very large job arrives and is used to create a left one-sided bin that creates a deficit of one job on the right, the next job from the same category will be used to create a right one-sided bin.

- If it is **a medium job**, a new one-sided bin is created with it. The same rules as for large and very large jobs are used to define the side where it will be placed. Moreover, when there is an unmatched medium job on one side, the next arriving medium job will be matched to this job on the opposite side to form a complete set of matched medium one-sided bins.

- If it is **a small job**, it is added to an incomplete one-sided bin of its job category if there is any. If there are multiple incomplete one-sided bins, the choice of the one-sided bin to which the new job is added is arbitrary and does not have an impact on the analysis derived in Section 4.4. If there are no incomplete one-sided bins, a new one-sided bin is opened with this new job as part of an incomplete set of matched one-sided bins or a new set of matched one-sided bins. The choice of the type of the one-sided bin which will be opened with this new job is arbitrary and does not have an impact on the analysis derived in Section 4.4.

The above rules imply that at any step in the algorithm, there is the same number of left and right large one-sided bins, up to a one job's difference. The same can be said for one-sided bins of medium jobs. The rules also imply that there can be at most one incomplete set of matched one-sided bins at a time for each job category.

4.3.3 Creating a Schedule

This section describes the process of how the complete sets of matched one-sided bins introduced above are converted in a schedule. This process can be applied at any step of the algorithm. Its goal is to rematch one-sided bins from complete sets of matched one-sided bins with large one-sided bins (jobs of length between $\frac{1}{2}$ and $\frac{2}{3}$) in order to obtain *complete sets of scheduled bins*. The obtained schedule depends on the complete sets of matched one-sided bins and large one-sided bins and therefore changes as they evolve. It should be noted that this process is reversible, and can therefore be implemented again after the arrival of a new job. In practice, the scheduling procedure would be applied only once, when all appointments for a specific day have been made.

Rematching Process.

The process described in this section creates *complete sets of rematched bins* from complete sets of matched one-sided bins. Consider first one-sided bins of small jobs. Note that two Type 2 fat one-sided bins can always be matched together because their respective scheduled time is smaller than $\frac{1}{2}$. Similarly, a Type 2 thin one-sided bin of small jobs can always be matched to a large one-sided bin because their respective scheduled times are smaller than $\frac{1}{3}$ and $\frac{2}{3}$. Recall that two Type 2 bins are said to be *symmetrical* if their Type 2 thin one-sided bins are placed on opposite sides. To perform rematching, two symmetrical Type 2 bins and two opposite (left and right) large one-sided bins are used. The Type 2 thin one-sided bins are rematched to the corresponding large one-sided bins on the opposite side, and the remaining Type 2 fat one-sided bins are rematched together. This is only possible because the selected Type 2 bins are symmetrical (fat/thin sides are on opposite sides). The rematching process therefore reduces the total number of bins.

Recall that n_2 is even. Therefore, in order to rematch a complete set of matched one-sided bins that contains n_2 Type 2 bins from a given job category, $\frac{n_2}{2}$ large one-sided bins must be selected on the left, and $\frac{n_2}{2}$ large one-sided bins on the right. If there is sufficient number of large one-sided bins to rematch the n_2 Type 2 bins from the set, the obtained set of bins is called a *complete set of rematched bins*. This is illustrated in Figure 4-5, where Figure 4-5(a) shows the initial complete set of matched one-sided bins and the selected large one-sided bins, and Figure 4-5(b) shows the complete set of rematched bins.

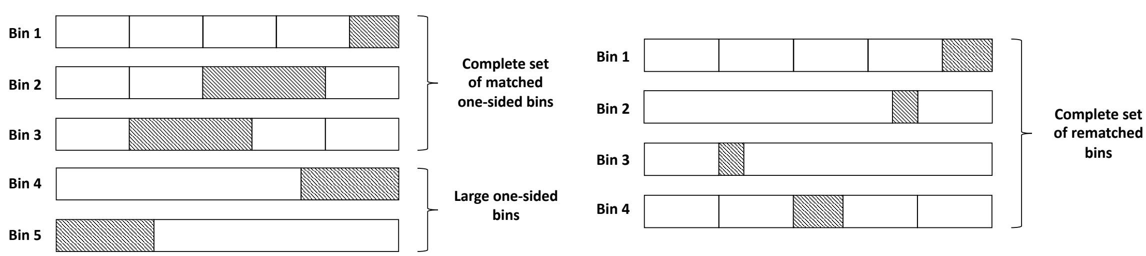

Figure 4-5: Rematching process on a complete set of matched one-sided bins of category $(\frac{1}{5}; \frac{1}{4}]$. From left to right: (a) Initial situation with the HR algorithm. In the complete set of matched one-sided bins, $n_1 = 1$ and $n_2 = 2$. Bin 1 corresponds to the Type 1 bin, while bin 2 and bin 3 are symmetrical Type 2 bins. Bin 4 and 5 are large one-sided bins. (b) Final situation after rematching with the HR algorithm. The large job from bin 4 has been rematched to the Type 2 thin one-sided bin from bin 2, and the large job from bin 5 has been rematched to the Type 2 thin one-sided bin from bin 3. The Type 2 fat one-sided bin from bin 2 has been rematched to the Type 2 fat one-sided bin from bin 3. Grey parts represent unoccupied space in the bin.

For medium jobs, the rematching process only consists of rematching each one-sided bin from a bin with a large one-sided bin from the opposite side as long as their total scheduled time does not exceed 1. It should be noted that this is not necessarily the case since medium jobs can have a scheduling time as large as $\frac{1}{2}$ and large jobs can have a scheduling time as large as $\frac{2}{3}$. This issue will be further addressed in Section 4.3.3.

Assignment of Additional Large Jobs Process.

A complete set of rematched bins contains a given number of large jobs. The process of *assignment of additional large jobs* includes the addition of large one-sided bins to a complete set of rematched bins, until the proportion of bins in the set containing a large job is larger than $\frac{2}{3}$. In the case where a complete set of rematched bins already has a proportion of bins in the set containing a large job larger than $\frac{2}{3}$, no additional large jobs need to be assigned to the set of bins.

The assignment process only means that the additional large one-sided bins will now be counted as bins in the new rematched set of bins where they have been assigned.

Complete Set of Scheduled Bins.

Consider a complete set of matched one-sided bins from some job category. A *complete set of scheduled bins* is created in the following manner:

1. A rematching process is applied to the complete set of matched one-sided bins as described in Section 4.3.3 and a *complete set of rematched bins* is obtained.

2. Additional large jobs are then assigned to the complete set of rematched bins as described in Section 4.3.3 to obtain a *complete set of scheduled bins* is obtained.

After the scheduling process, any initial set of bins where step 1 or step 2 could not be completed totally is called a *complete set of partially scheduled bins.* Large one-sided bins that have not been rematched or assigned to any complete set of bins are called *unassigned large one-sided bins.*

Summary of Scheduling Process.

The scheduling process takes as input complete sets of matched one-sided bins and unassigned large one-sided bins. It will be shown that there are only finite number of incomplete sets of matched one-sided bins, and therefore they do not impact the asymptotic analysis and therefore will not be considered in this section.

The following steps will be executed in order as long as there are still unassigned large one-sided bins.

Step 1 : Large jobs on the right and on the left are ordered by decreasing size, and are used in that order for rematching and assignement process. Complete sets of matched one-sided bins of small jobs are then selected by decreasing job category to form complete sets of scheduled bins according to the process described in Section 4.3.3. This is illustrated in Figure 4-6(a). This first step of the scheduled process can end in two situations. The first situation is when all complete sets of matched one-sided bins were scheduled to form complete sets of scheduled bins. There still potentially remains some unassigned large one-sided bins. The second situation is when there are no additional unassigned large one-sided bins. In that case, there remains at most one set of partially scheduled bins where either the rematching or the large jobs assignment process could not be completed. There also remains some complete sets of matched one-sided bins.

Step 2 : If the first step stops in the first situation, remaining unassigned large one-sided bins are used to create complete sets of scheduled bins from complete sets of one-sided bins of medium jobs. Large one-sided bins are now ordered by *increasing size* on the left and on the right, and medium one-sided bins are ordered by *increasing size* on the left and on the right, and they are both used in that order. Pairs of large one-sided bins and medium one-sided bins are then selected to form complete sets of scheduled bins as described at the end of Section 4.3.3, as long as they meet the scheduling time requirement. This is illustrated in Figure 4-6(b).

The scheduling process of medium jobs can stop in the following three situations:

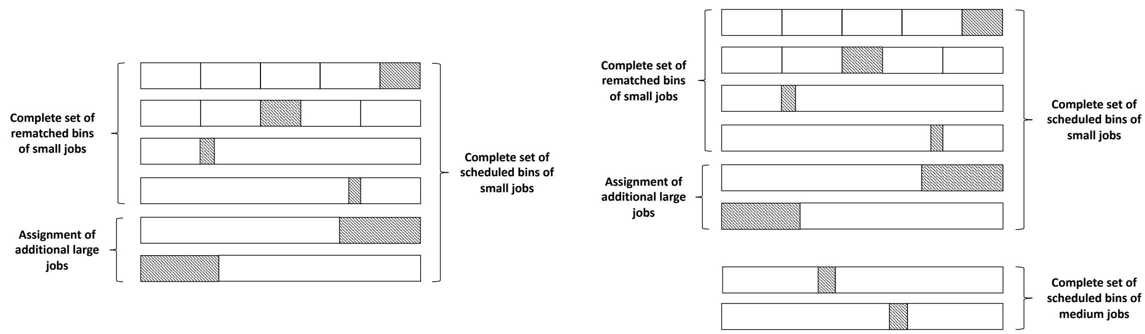

Figure 4-6: Scheduling Process. From left to right: (a) First step of the scheduling process (only one set of matched one-sided bins of small jobs is depicted for sake of simplicity). Large one-sided bins are ordered from largest to shortest on both sides, and are rematched and assigned to complete sets of matched one-sided bins of small jobs to create complete sets of scheduled bins. (b) Second step of the scheduling process. Large one-sided bins are ordered from shortest to largest on both sides, and are rematched to complete sets of matched one-sided bins of medium jobs to create complete sets of scheduled bins of medium jobs.

- All complete set of matched one-sided bins of medium jobs have been rematched. In that case, there potentially remains some unassigned large one-sided bins.

- All remaining unassigned large one-sided bins have been used in the rematching process of medium jobs. In that case, there potentially remains some complete sets of partially scheduled bins of medium jobs.

- Since there is no guarantee that any large and medium one-sided bin can be matched (the sum of their sizes could exceed 1), it is possible that there remains both unassigned large one-sided bins and complete sets of partially scheduled bins of medium jobs at the end of the scheduling process. This last case is illustrated in Figure 4-7.

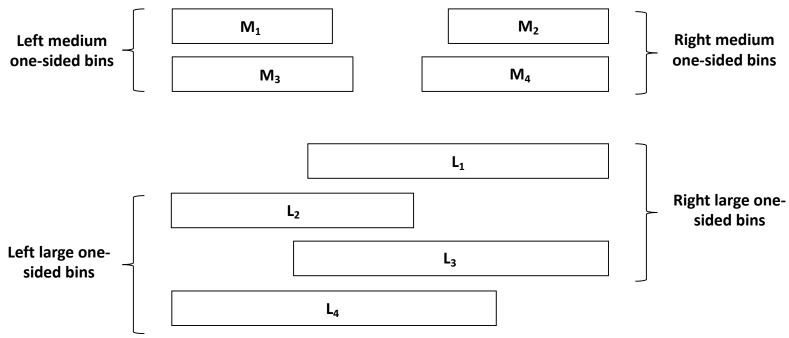

Figure 4-7: Stopping criteria in step 2 of the scheduling process. Large and medium jobs are ordered by increasing size on each side. Here, medium job M_1 cannot fit with large job L_1, and the ordering of jobs by increasing length imply that medium job M_3 cannot fit with neither large job L_1 or L_3. Therefore, even if medium job M_2 can fit with large job L_2, the rematching process stops here.

4.4 Competitive Analysis of the Online Scheduling Algorithm

4.4.1 Overview of the Analysis

The goal of the competitive analysis is to show that as the number of jobs goes to infinity, the HR algorithm obtains an asymptotic competitive ratio of $\frac{3}{2}$.

The analysis will be organized as follows. It starts with establishing two essential properties of complete sets of scheduled bins regarding fullness and proportion of large jobs. Specifically, it will be shown in Lemma 6 that all complete sets of scheduled bins are at least $\frac{2}{3}$ full and have a proportion of bins containing a large job greater or equal than $\frac{2}{3}$. Secondly, the analysis shows in Lemma 7 a property of the size distribution of large and medium jobs on each side. Finally, three different scenarios that can be encountered during the scheduling process are analyzed. The asymptotic competitive ratio will be derived for the first two scenarios in Lemma 8 and 9 by using two lower bounds. The first bound comes from fullness of bins. If all bins in the schedule are at least $\frac{2}{3}$ full, then OPT must use at least $\frac{2}{3}$ of the bins used by the HR algorithm to fit all jobs. The asymptotic ratio of $\frac{3}{2}$ directly follows. The second bound comes from the proportion of large jobs in the bins. If the proportion of bins containing a large job in the schedule is at least $\frac{2}{3}$, then OPT must again use at least $\frac{2}{3}$ of the bins used by the HR algorithm to fit all jobs since it can pack at most one large job per bin. The asymptotic ratio of $\frac{3}{2}$ directly follows. The asymptotic ratio for the last scenario considered in Lemma 10 will be derived by focusing on the size distribution of large and medium jobs on each side.

4.4.2 Properties of Complete Sets of Scheduled Bins

It will first be shown in Lemma 1 that all complete sets of matched one-sided bins are at least $\frac{2}{3}$ full.

Lemma 1. *For each job category, the average fullness of a complete set of matched one-sided bins is greater or equal than $\frac{2}{3}$.*

Proof. For categories of small jobs except category $[0; \frac{1}{12}]$, the proportion of Type 1 and Type 2 fat and thin one-sided bins in a complete set of matched one-sided bins is fixed by Appendix D, as well as the number of jobs in each one-sided bin . Knowing the minimum job's length for each category, this directly yields the fullness result for each job category separately. For medium jobs, the result follows from the fact that each one-sided bin contains a job larger than $\frac{1}{3}$. Since complete set of matched medium one-sided bins contains only one bin with two one-sided bins, the lemma holds for this category as well. For category $[0; \frac{1}{12}]$, the Type 2 thin one-sided bin is larger than $\frac{1}{3} - \frac{1}{12}$ and the Type 2 fat one-sided bin is larger than $\frac{1}{2} - \frac{1}{12}$ as specified in Appendix D. Therefore, the total size of a Type 2 bin from a complete set of matched one-sided bins must be larger than $\frac{1}{3} + \frac{1}{2} - \frac{2}{12} = \frac{2}{3}$. Since complete sets of matched one-sided bins are only made of Type 2 bins for this job category, the lemma also holds. □

It will next be shown that complete set of rematched bins are also at least $\frac{2}{3}$ full. To that end, Lemma 2 shows that any partial rematching process implemented on a set of bins with average fullness greater or equal than $\frac{2}{3}$ maintains an average fullness greater or equal than $\frac{2}{3}$.

Lemma 2. *Consider a complete set of matched one-sided bins from some job category with n_1 Type 1 bins and n_2 Type 2 bins and with a minimum average fullness greater or equal than $\frac{2}{3}$. Performing partial rematching between two large one-sided bins and two symmetrical Type 2 bins from the complete set of bins creates a new set of bins with average fullness greater or equal than $\frac{2}{3}$.*

Proof. As described in Section 4.3.3, two large one-sided bins are used to rematch two Type 2 bins. Three bins are obtained from this rematching process, made of the combination of the large one-sided bins with the Type 2 thin one-sided bins, and of the combination of the Type 2 fat one-sided bins. Let f_1, f_2, f_3 be the respective minimum fullness level of Type 1 bins, Type 2 thin one-sided bins and Type 2 fat one-sided bins for the same job category. By hypothesis, the average fullness f is such that $f \geq \frac{n_1 f_1 + n_2(f_2+f_3)}{n_1+n_2} \geq \frac{2}{3}$. Because large jobs are at least $\frac{1}{2}$ full, the average fullness of the new set of bins verifies:

$$\begin{aligned}
\tilde{f} &\geq \frac{n_1 f_1 + (n_2 - 2)(f_2 + f_3) + (1 + 2f_2 + 2f_3)}{n_1 + n_2 + 1} \\
&\geq \frac{n_1 f_1 + n_2(f_2 + f_3)}{n_1 + n_2 + 1} + \frac{1}{n_1 + n_2 + 1} \\
&\geq \frac{2}{3}\frac{n_1 + n_2}{n_1 + n_2 + 1} + \frac{1}{n_1 + n_2 + 1} \\
&\geq \frac{2}{3}\frac{n_1 + n_2 + 1}{n_1 + n_2 + 1} \\
&\geq \frac{2}{3}.
\end{aligned}$$

□

Finally, Lemma 3 shows that all complete sets of rematched bins have an average fullness greater or equal than $\frac{2}{3}$.

Lemma 3. *For each job category, the average fullness of a complete set of rematched bins is greater or equal than $\frac{2}{3}$.*

Proof. Complete sets of matched one-sided bins are at least $\frac{2}{3}$ full by Lemma 1. Lemma 2 shows that if a complete set matched one-sided bins is at least $\frac{2}{3}$ full, it will remain at least $\frac{2}{3}$ full after the implementation of any partial rematching process. A complete set of rematched bins of small jobs is obtained by applying a rematching process on a complete set of matched one-sided bins that is at least $\frac{2}{3}$ full. Therefore, a rematched set of bins of small jobs will remain at least $\frac{2}{3}$ full. For medium jobs, the result directly holds since a rematched set of bins contains 2 bins, each bin made of one large job at least $\frac{1}{2}$ full and one medium job at least $\frac{1}{3}$ full. $\square$

It will next be shown that for each job category, additional large one-sided bins can be assigned to a complete set of rematched bins in order to reach a proportion of bins containing a large job greater or equal than $\frac{2}{3}$, while ensuring that the new set of bins maintains an average fullness larger than $\frac{2}{3}$.

Lemma 4. *For a complete set of rematched bins of category* $(\frac{1}{j+1}; \frac{1}{j}]$ *for* $j = 6, 8, 9, 10, 11$, *or* $[0; \frac{1}{12}]$, *or of medium jobs, the proportion of bins containing a large job is greater or equal than* $\frac{2}{3}$.

Proof. For categories $(\frac{1}{j+1}; \frac{1}{j}]$ for $j = 6, 8, 9, 10, 11$ or category $[0; \frac{1}{12}]$, rematched sets of bins are made of 3 bins. Two of them are made of a Type 2 thin one-sided bin rematched to a large one-sided bin, and one of them is made of two Type 2 fat one-sided bins matched together. This directly yields that the proportion of bins containing a large job is equal to $\frac{2}{3}$. For medium jobs, a rematched set of bins is made of 2 bins, each one of them containing a large job. The proportion of bins containing a large job is equal to 1. $\square$

Lemma 5. *For a complete set of rematched bins of category* $(\frac{1}{j+1}; \frac{1}{j}]$ *for* $j = 3, 4, 5, 7$, *additional large one-sided bins can be assigned to a complete set of rematched bins so that the proportion of bins containing a large job in the new set of bins is larger than* $\frac{2}{3}$, *and the average fullness of the new set of bins is greater or equal than* $\frac{2}{3}$.

Proof. Let's prove the lemma for each job category separately:

1. For $j = 3$: There are 7 bins in a complete set of rematched bins from this job category. Four are Type 1 bins with minimum fullness of $\frac{3}{4}$, two are obtained from Type 2 thin one-sided bins rematched with large one-sided bins and with

minimum fullness $\frac{3}{4}$, and one is obtained from two rematched Type 2 fat one-sided bins and with minimum fullness $\frac{1}{2}$. Each one-sided bin from this last bin can be rematched with an additional large one-sided bin. After performing this rematching, the new set of bins is made of eight bins, each of them at least $\frac{3}{4}$ full. This means that 4 large jobs can still be assigned to this set of bins to form a new set of bins while ensuring a minimum average fullness superior to $\frac{2}{3}$. The new proportion of bins containing a large job in this set of bins is equal to $\frac{2}{3}$. This is illustrated in Figure 4-8.

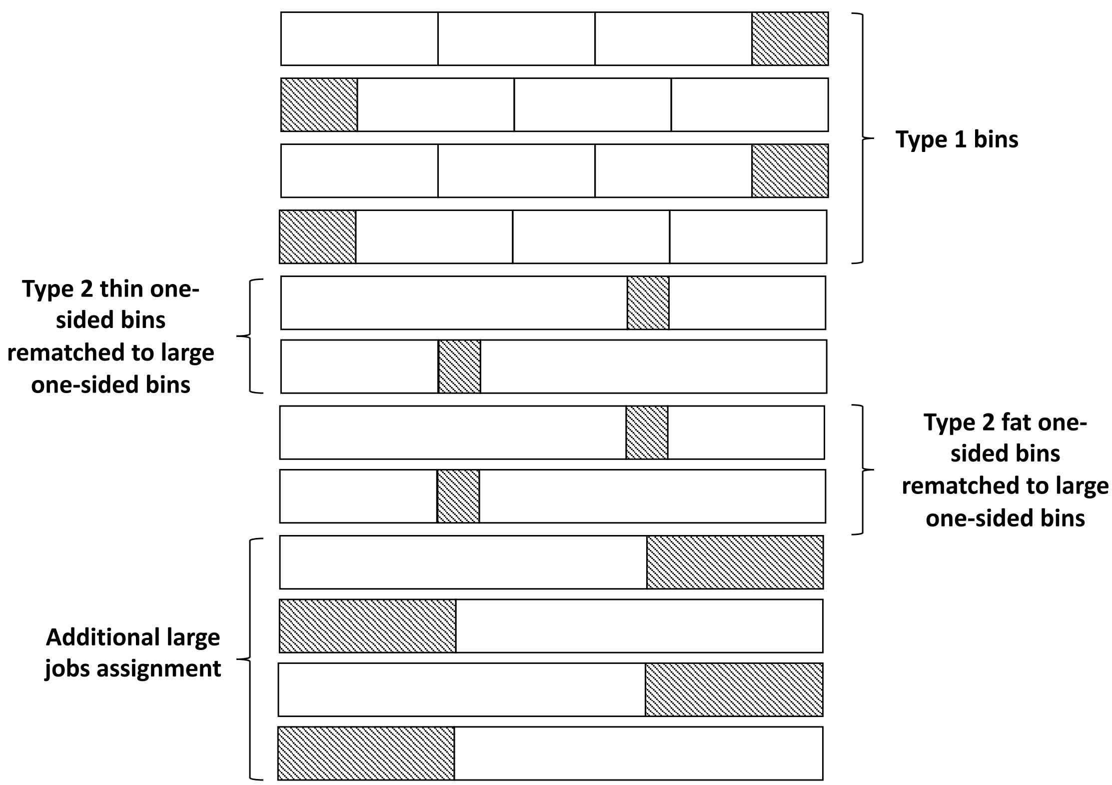

Figure 4-8: Proof of Lemma 5 for Job Category $(\frac{1}{4}; \frac{1}{3}]$

2. For $j = 4$: There are 4 bins in a complete set of rematched bins from this job category. One is a Type 1 bin with minimum fullness $\frac{4}{5}$, two are obtained from rematched Type 2 thin one-sided bins with large one-sided bins and with minimum fullness $\frac{7}{10}$, and one is obtained from two rematched Type 2 fat one-sided bins and with minimum fullness $\frac{4}{5}$. 2 large jobs can be assigned to this set

of rematched bins, while maintaining a minimum average fullness superior to $\frac{2}{3}$. In this new set of rematched bins made of 6 bins, 4 of them contain a large job, which implies that the new proportion of large jobs is equal to $\frac{2}{3}$. This is illustrated in Figure 4-9.

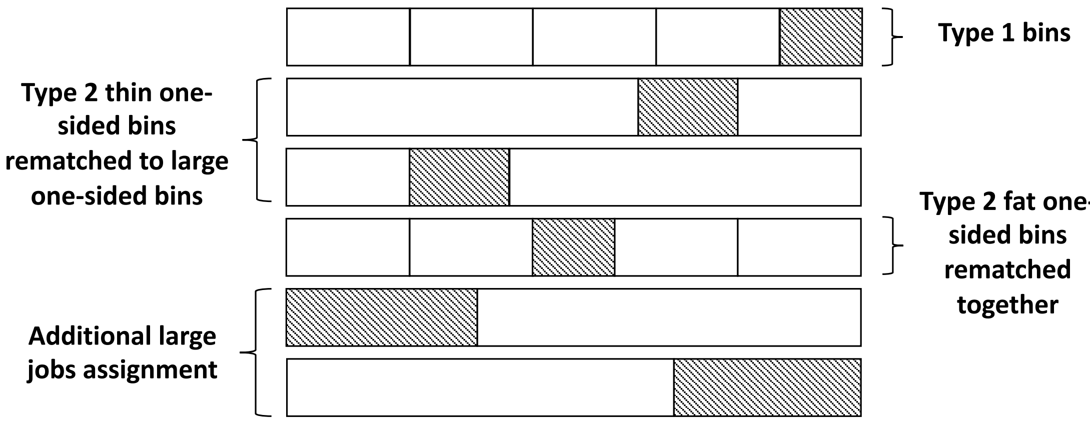

Figure 4-9: Proof of Lemma 5 for Job Category $(\frac{1}{5}; \frac{1}{4}]$

3. For $j = 5$: Let's first make the following remark. A one-sided bin made of two jobs from this category can always be matched with a large job of size $(\frac{1}{2}; \frac{3}{5}]$. However, such a one-sided bin cannot necessarily be rematched with a large job of size $(\frac{3}{5}; \frac{2}{3}]$. Now, in a complete set of rematched bins of this job category, there are 5 bins: two are Type 1 bins with fullness $\frac{5}{6}$, two are obtained from rematched Type 2 thin one-sided bins with large one-sided bins and with minimum fullness $\frac{2}{3}$, and one is obtained from two rematched Type 2 fat one-sided bins and with minimum fullness $\frac{2}{3}$.

   Let's assume first that there are two available large jobs of size included in $(\frac{1}{2}, \frac{3}{5}]$. Those two additional large jobs can be rematched to the Type 2 fat one-sided bins from the set of rematched bins. In this new set of rematched bins made of 6 bins, 4 of them contain a large job, which implies that the new proportion of large jobs is equal to $\frac{2}{3}$. The minimum average fullness is superior to $\frac{2}{3}$. This is illustrated in Figure 4-10.

Let's now assume that there are no two large jobs of size included in $(\frac{1}{2}, \frac{3}{5}]$ available. Since what matters in the analysis is the asymptotic competitive ratio, it can be assumed that there are at least four available large jobs of size included in $(\frac{3}{5}, \frac{2}{3}]$ (otherwise only a finite number of jobs are concerned). Four such large jobs can then be assigned to the set of rematched bins while maintaining a minimum average fullness superior to $\frac{2}{3}$. In this new set of rematched bins made of nine bins, six of them contain a large job, which implies that the new proportion of large jobs is equal to $\frac{2}{3}$. Thi is illustrated in Figure 4-11.

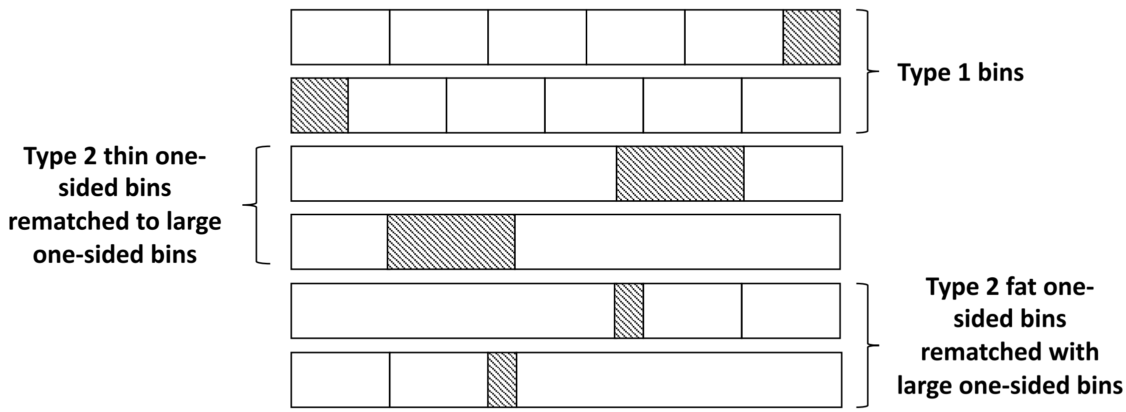

Figure 4-10: First Scenario of Proof of Lemma 5 for Job Category $(\frac{1}{6}; \frac{1}{5}]$

4. For $j = 7$: There are 17 bins in a complete set of rematched bins from this job category. Two are Type 1 bins with minimum fullness $\frac{7}{8}$, ten are obtained from rematched Type 2 thin one-sided bins with one-sided bins of large jobs and with minimum fullness $\frac{3}{4}$, and five are obtained from rematched Type 2 fat one-sided bins and with minimum fullness $\frac{3}{4}$. Therefore, 4 large jobs can be added to this set while ensuring a minimum average fullness superior to $\frac{2}{3}$. In this new set of rematched bins made of 21 bins, 14 of them contain a large job, which implies that the new proportion of large jobs is equal to $\frac{2}{3}$.

$\square$

The next lemma follows from Lemma 4 and Lemma 5.

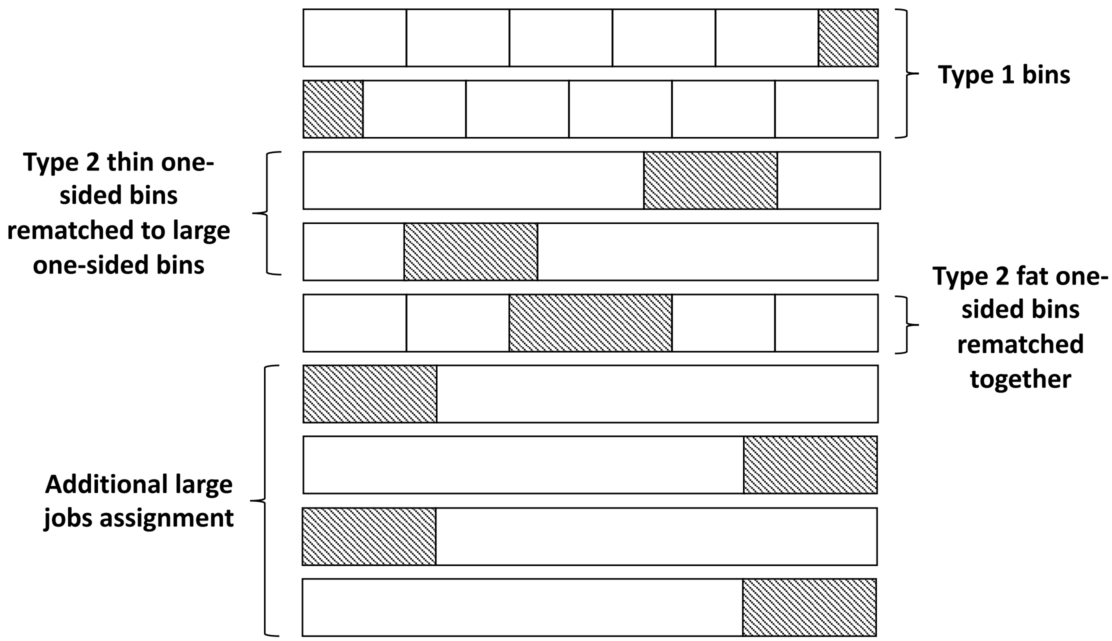

Figure 4-11: Second Scenario of Proof of Lemma 5 for Job Category $(\frac{1}{6}; \frac{1}{5}]$

Lemma 6. *A complete set of scheduled bins is at least $\frac{2}{3}$ full, and has a proportion of bins containing a large job greater or equal than $\frac{2}{3}$.*

Proof. Lemmas 4 and 5 show that large one-sided bins can be assigned to a complete set of rematched bins from any job category so that the new set of bins has a proportion of bins containing a large job at least equal to $\frac{2}{3}$, and an average fullness greater or equal then $\frac{2}{3}$. This proves that the additional large job assignment process described in Section 4.3.3 can always be applied while maintaining the average fullness of the new set of bins above $\frac{2}{3}$. Therefore, a complete set of scheduled bins presents the two properties of fullness and proportion of large jobs. □

4.4.3 Property of Bounded Imbalance for Large and Medium Jobs

This section is concerned with another property of the algorithm related to the notion of *imbalance* for large and medium jobs. It is obvious that if one considers the jobs in the order they arrived, the algorithm is balanced with respect to the number of large jobs on the left side and on the right side up to a constant of 1 for each iteration. However, it is not necessarily true that this holds when large jobs are ordered by increasing size.

Consider a sequence of i large jobs, ordered from shortest to longest. Let $Y_k = 1$ denote the event that the k^{th} large job is placed on the left, and $Y_k = -1$ the event that it is placed on the right. Let I_i denote the imbalance of the i^{th} smallest large jobs. Then

$$I_i = \sum_{k=1}^{i} Y_k.$$

If $I_i > 0$, there are more of the i^{th} smallest large jobs on the left side than on the right side, and if $I_i < 0$ there are more of the i^{th} smallest large jobs on the right side than on the left side.

The following lemma shows that for any ordered subset of large jobs from a sequence of L large jobs, the imbalance can be uniformly bounded.

Lemma 7. *Consider L large jobs ordered from shortest to longest. Let $(a_1, \ldots, a_L)$ denote their respective arrival time where $1 \leq a_k \leq L$, and $(Y_1, \ldots, Y_L)$ the corresponding side they are matched to upon arrival, where $Y_k = 1$ if the k^{th} smallest large job is placed on the left and $Y_k = -1$ if it is placed on the right. Now, consider the i^{th} first smallest large jobs in this sequence. Then, there exists a constant C independent of L and i such that with probability 1,*

$$|I_i| \leq C\sqrt{L \log \log L} \quad \forall i.$$

Proof. While the large jobs are ordered by size, they did not arrive necessarily in this order. Consider all the pairs of jobs among the first i smallest large jobs that arrived consecutively, ie with arrival times are of the form $(a_k, a_l) = (2n-1, 2n)$ with $n \in \mathbb{N}$. By the matching rules of the HR algorithm, those jobs will be placed in an alternating fashion on the left side and the right side which implies that either $Y_k = 1$ and $Y_l = -1$, or $Y_k = -1$ and $Y_l = 1$. Therefore, such pairs of jobs will not contribute to the imbalance. Let's now consider the worst-case scenario in terms of imbalance, where there are no such pairs in the sequence $(a_1, \ldots, a_i)$. This implies that no jobs in the sequence of the i^{th} smallest large jobs arrived in a consecutive manner. However, this also implies by the randomization process used to match large jobs that the random variables $(Y_1, \ldots, Y_i)$ are all independent, following a distribution such that $P(Y_k = 1) = \frac{1}{2}$ and $P(Y_k = -1) = \frac{1}{2}$. Recall that the corresponding imbalance is $I_i = \sum_{k=1}^{i} Y_k$. By the law of iterated logarithm [81], since $\mathbb{E}[Y_k] = 0$ and $\mathbb{V}ar[Y_k] = 1$,

$$\limsup_{i \to \infty} \frac{\pm I_i}{\sqrt{2i \log \log i}} = 1 \quad \text{a.s.}$$

Therefore, there exists a constant C independent of i such that for all i almost surely, $|I_i| \leq C\sqrt{i \log \log i}$. Finally, since $i \leq L$, $|I_i| \leq C\sqrt{L \log \log L}$.

□

Analogous results are implied by Lemma 7 for medium jobs.

4.4.4 An Asymptotic Competitive Ratio of $3/2$

Notations.

Consider the following notations:

- Let v denote the number of one-sided bins with very large jobs (these one-sided bins actually form a single bin in the final schedule).
- Let z denote the number of bins in complete sets of scheduled bins (from either small or medium jobs).
- Let n denote the number of bins from complete sets of partially scheduled bins of small jobs.
- Let m denote the number of one-sided bins from complete sets of partially scheduled bins of medium jobs.
- Let ℓ denote the number of unassigned large one-sided bins.

- Let k denote the number of bins included in incomplete sets of matched one-sided bins.

First, note that k is bounded by a given constant K independent of the number of jobs. Specifically, there can be at most one incomplete set of matched one-sided bins per job category. If n_j is the number of bins in a complete set of matched one-sided bins of the job category j, this implies that there can be at most $n_j - 1$ bins included in an incomplete set of matched one-sided bins for job category j. If N is the total number of categories of jobs, this implies that there can be at most $K = \sum_{j=1}^{N}(n_j - 1)$ bins included in incomplete sets of bins. Therefore, $k \leq K$, where K is a constant independent on the number of jobs in the schedule.

Based on previous section, the analysis considers four scenarios that can be encountered after the scheduling process terminates.

1. There are complete sets of partially scheduled bins of small jobs. This implies that there are no unassigned large one-sided bins as shown in Section 4.3.3 because of the scheduling process. Therefore, $\ell = 0$ and $n > 0$.

2. There are no complete sets of partially scheduled bins of small jobs, but there are partially complete set of scheduled bins of medium jobs and no unassigned large one-sided bins. Therefore, $n = 0$, $\ell = 0$ and $m > 0$.

3. There are unassigned large one-sided bins and only complete sets of scheduled bins of small and medium jobs. Therefore, $\ell > 0$, $n = 0$ and $m = 0$

4. There are unassigned large one-sided bins and complete sets of partially scheduled bins of medium jobs. Therefore, $\ell > 0$ and $m > 0$. This is a pathological scenario, which can lead to having a lot of large one-sided bins, but also complete sets of partially scheduled bins of medium jobs which do not have a proportion of bins containing a large job larger that $\frac{2}{3}$.

The first and second scenarios can actually be regarded as the same scenario, since this simply corresponds to a setting where there are no unassigned large one-sided bins.

Scenario 1: No Unassigned Large One-Sided Bins.

Lemma 8. *If there are no unassigned large one-sided bins ($\ell = 0$), then $ALG \leq \frac{3}{2}OPT + K$ where K is independent on OPT.*

Proof. The fact that there are no unassigned large one-sided bins implies that they have all been used in the rematching process and the assignment process. All complete sets of scheduled bins have an average fullness larger than $\frac{2}{3}$ by Lemma 6, so that all z bins are on average at least $\frac{2}{3}$ full. All complete sets of matched one-sided bins are also on average at least $\frac{2}{3}$ full by Lemma 1. Finally, all complete sets of partially scheduled bins have an average fullness larger than $\frac{2}{3}$ by Lemma 2 and Lemma 3. This implies that all n and m bins are also on average at least $\frac{2}{3}$ full. Finally, all very large jobs are at least $\frac{2}{3}$ full. This implies that $OPT \geq \frac{2}{3}(z + v + n + m)$. Moreover, $ALG = v + z + n + m + k$ and $k \leq K$ where K is independent on z, n, m and v. Therefore,

$$\begin{aligned}\frac{ALG}{OPT} &\leq \frac{v+z+n+m}{\frac{2}{3}(v+z+n+m)} + \frac{K}{OPT}\\ &\leq \frac{3}{2} + \frac{K}{OPT}.\end{aligned}$$

Therefore, $ALG \leq \frac{3}{2}OPT + K$.

□

Scenario 2: Unassigned Large One-Sided Bins and Complete Set of Scheduled Bins.

Lemma 9. *If there are unassigned large one-sided bins and only complete sets of scheduled bins ($n = 0$, $m = 0$ and $\ell > 0$), then $ALG \leq \frac{3}{2}OPT + K$ where K is independent on OPT.*

Proof. From Lemma 6, complete sets of scheduled bins have a proportion of bins containing large jobs greater or equal than $\frac{2}{3}$. Therefore, the z bins contain at least $\frac{2}{3}z$ large jobs. By hypothesis, $n = 0$ and $m = 0$ since all sets of bins are complete sets of scheduled bins. This implies that $OPT \geq v + \frac{2}{3}z + \ell$, since OPT cannot pack more than one large job per bin. Moreover, $ALG = v + z + \ell + k$ and $k \leq K$ which implies that $ALG \leq v + z + \ell + K$. Therefore,

$$\begin{aligned} \frac{ALG}{OPT} &\leq \frac{v + z + \ell}{\frac{2}{3}z + v + \ell} + \frac{K}{OPT} \\ &\leq \frac{3}{2} + \frac{K}{OPT}. \end{aligned}$$

And thus, $ALG \leq \frac{3}{2}OPT + K$.

□

Scenario 3: Unassigned Large One-Sided Bins and Partially Complete Set of Scheduled Bins of Medium Jobs.

As stated at the end of Section 4.3.3, this scenario occurs when some of the remaining unassigned large jobs cannot match with opposite medium one-sided bins from sets of partially scheduled bins. Recall that the total number of large one-sided bins is denoted by L, and the total number of medium one-sided bins is denoted by M. After having created all the complete sets of scheduled bins of small jobs, order the remaining one-sided bins of large jobs from shortest to largest on the right and left side as described in Section 4.3.3. The *overall rank* of a large job is the job's rank among all large jobs ordered by increasing size. Similarly, the *overall rank* of a medium job is the job's rank among all medium jobs ordered by increasing size.

Let L' be the overall rank of the first large one-sided bin that cannot be assigned anymore to the corresponding medium job on the opposite side. Without loss of generality, consider it is located on the right side. Figure 4-12 illustrates an example in which $L' = 6$. Let M' be the overall rank of the corresponding medium one-sided bin on the left. In Figure 4-12, $M' = 5$.

All medium one-sided bins on the left with overall rank higher than M' cannot be matched with any large one-sided bins on the right with overall rank higher than L', since large jobs and medium jobs are ordered by increasing size on each side. However, the same cannot necessarily be said regarding medium one-sided bins on the right and large one-sided bins on the left. In the worst case, it could be that all large one-sided bins on the left could match with all medium one-sided bins on the right. For example, on Figure 4-12, medium job with overall rank equal to 3 on the right can be matched with large job with overall rank 3 on the left.

Let $I_{L'}$ and $I_{M'}$ be the respective imbalance for the sequence of the shortest $L'-1$ large jobs and the sequence of the shortest $M'-1$ medium jobs, as defined in Section 4.4.3. Recall that Y_i is the random variable taking value 1 if the i^{th} large job is placed on the left, and -1 otherwise. Define similarly Z_i the random variable taking value 1 if the i^{th} medium job is placed on the left, and -1 otherwise. Then, $I_{L'} = \sum_{k=1}^{L'-1} Y_k$ and $I_{M'} = \sum_{k=1}^{M'-1} Z_k$.

It should be noted that the imbalance for medium jobs is not necessarily equal to the imbalance for large jobs as is shown on Figure 4-12 where $I_{L'} = 3$ while $I_{M'} = 2$. Lemma 7 implies that $I_{L'}$ and $I_{M'}$ can be bounded and this will play a major role in the subsequent analysis.

Lemma 10. *Consider an infinite sequence of iterations where there are unassigned large one-sided bins, and complete sets of partially scheduled bins of medium jobs (ie $\ell > 0$ and $m > 0$). Then almost surely,*

$$\lim_{N\to\infty} \sup_{I:OPT(I)\geq N} \frac{ALG(I)}{OPT(I)} \leq \frac{3}{2}.$$

Proof. Recall that L and M denote the respective total number of large and medium one-sided bins. For the sake of clarity, the index i will be omitted when considering quantities L, M, ℓ or m for a given iteration. If either $\lim_{i\to\infty} L < \infty$ or $\lim_{i\to\infty} M < \infty$ (i.e., finite number of large or medium jobs), then scenario 3 is actually equivalent to scenario 1 or 2, since the number of unassigned large jobs or the number of partially complete scheduled set of medium bins is finite. Therefore, it will now be considered that $\lim_{i\to\infty} L = \infty$ and $\lim_{i\to\infty} M = \infty$.

Consider first the case where both $I_{L'} = 0$ and $I_{M'} = 0$. Then all ℓ large jobs cannot match with any of the m medium job since large and medium jobs are ordered by increasing size on the left side and on the right side. Therefore, two unassigned

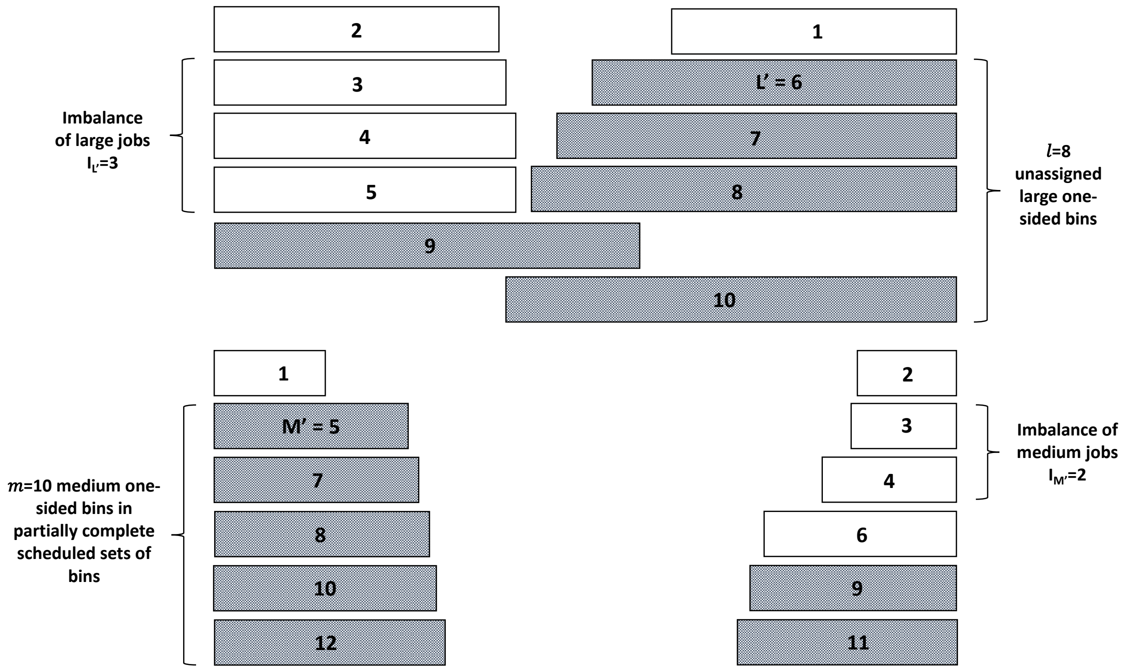

Figure 4-12: Scenario with Unassigned Large Jobs. Large and medium jobs are ordered by increasing size from top to bottom. Large jobs are represented at the top, while medium jobs are represented at the bottom. There are $\ell = 8$ unassigned large one-sided bins, and $m = 10$ medium one-sided bins in complete sets of partially scheduled bins. Numbers on the bins represent each job's overall rank with respect to its job category. The first large job that cannot be matched to a medium job anymore is located on the right, with overall rank $L' = 6$. The corresponding medium job on the left has overall rank $M' = 5$. The grey colouring represents jobs that cannot be matched to the corresponding large or medium job located on the other side. The large jobs imbalance is equal to $I_{L'} = 3$, while the medium jobs imbalance is equal to $I_{M'} = 2$.

large jobs can be assigned to a complete set of partially scheduled medium bins to form a new set of bin where average fullness is higher than $\frac{2}{3}$ while the proportion of bins containing a large job is $\frac{2}{3}$. This is illustrated in Figure 4-13. This new set of bins verifies the two conditions of fullness and proportion of large jobs defining a complete set of scheduled bins, as specified in Section 4.3.3. Therefore, the scenario 3 is now equivalent to scenario 1 if $m > \ell$, and to scenario 2 if $\ell \geq m$. By Lemma 8 and Lemma 9, this implies that

$$\lim_{N\to\infty} \sup_{I:OPT(I)\geq N} \frac{ALG(I)}{OPT(I)} \leq \frac{3}{2}.$$

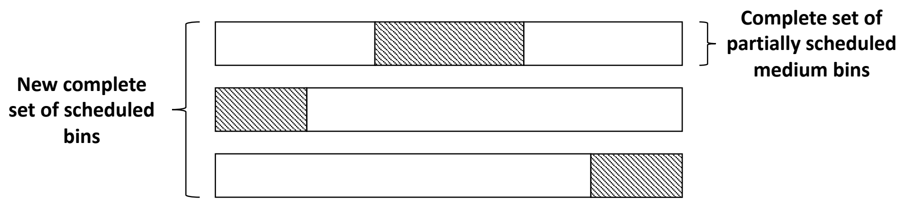

Figure 4-13: Scenario Without Imbalance. New set of bins obtained by assigning two additional large jobs to the initial complete set of partially scheduled medium bins. Since both large jobs cannot match with the medium job on the opposite side, the minimum fullness of this new set of bins is at least $\frac{2}{3}$. Moreover, the proportion of bins containing a large job in this new set is equal to $\frac{2}{3}$. Therefore, this new set of bins can be called a complete set of scheduled bins.

Consider now the case where there is imbalance either in large or medium jobs. Denote by $\bar{I} = \max\{I_{L'}, I_{M'}\}$ the maximum imbalance. Any pair of large jobs where both jobs have overall rank higher than L' can still be assigned to any pair of medium jobs where both jobs have overall rank higher than M' to form a complete set of scheduled bins as shown in Figure 4-13. However, the same cannot be said for pairs of large and medium jobs where one of the job has overall rank lower than L' or M' respectively. Therefore, there are $\bar{I}$ pairs of large jobs and $\bar{I}$ pairs of medium jobs where the bounds on fullness and proportion of large jobs cannot be derived. It will now be shown that such pairs can actually be ignored in the analysis.

As shown in Lemma 7, it is true that for all L' and M' almost surely,

$$I_{L'} \leq C\sqrt{L \log \log L} \quad \text{and} \quad I_{M'} \leq C\sqrt{M \log \log M}.$$

Moreover, $OPT \geq L$ and $OPT \geq \frac{M}{2}$ since OPT can pack at most one large job per bin and two medium jobs per bin. Therefore, almost surely,

$$\begin{aligned}\lim_{N\to\infty} \sup_{I:OPT(I)\geq N} \frac{\bar{I}}{OPT} &\leq \lim_{N\to\infty} \sup_{I:OPT(I)\geq N} \max\left\{\frac{C\sqrt{L\log\log L}}{L}, \frac{2C\sqrt{M\log\log M}}{M}\right\} \\ &\leq \max\left\{\lim_{L\to\infty} \frac{C\sqrt{L\log\log L}}{L}, \lim_{M\to\infty} \frac{2C\sqrt{M\log\log M}}{M}\right\}.\end{aligned}$$

Therefore, almost surely,

$$\lim_{N\to\infty} \sup_{I:OPT(I)\geq N} \frac{\bar{I}}{OPT} = 0.$$

This implies that just like the k incomplete sets of bins, the $\bar{I}$ pairs of large and medium jobs can be disregarded in the asymptotic analysis of the ratio $\frac{ALG}{OPT}$. Therefore, scenario 3 is again equivalent in terms of asymptotic behavior to scenario 1 if $m > \ell$, and to scenario 2 if $\ell \geq m$. By Lemma 8 and Lemma 9, this implies that $\lim_{N\to\infty} \sup_{I:OPT(I)\geq N} \frac{ALG(I)}{OPT(I)} \leq \frac{3}{2}$.

□

Combining previous sections, we can therefore obtain the following main result:

Theorem 1. *Consider an infinite sequence of job arrivals. The HR algorithm has an asymptotic competitive ratio of $\frac{3}{2}$ almost surely:*

$$\lim_{N\to\infty} \sup_{I:OPT(I)\geq N} \frac{ALG(I)}{OPT(I)} \leq \frac{3}{2}.$$

Proof. Consider three subsequences of iterations, one for each of the three scenarios specified previously. Lemma 8, 9 and 10 respectively prove that each infinite subsequence verifies that almost surely $\lim_{N\to\infty} \sup_{I:OPT(I)\geq N} \frac{ALG(I)}{OPT(I)} \leq \frac{3}{2}$. Therefore, combining those three results, it is directly obtained that, with probability 1, $\lim_{N\to\infty} \sup_{I:OPT(I)\geq N} \frac{ALG(I)}{OPT(I)} \leq \frac{3}{2}$. □

4.5 Computational Experiments

In this section, the performance of the HR algorithm is evaluated over randomly generated sequences of appointment requests. The resulting objective values are compared to the respective optimal offline objective values. The analyses also examines how the performance of the HR algorithm changes with respect to the number of appointments n. For each value of n, 1,000 independent sequences of appointments (each of size n) are generated. Within each sequence of appointments, the duration of the n appointments are independent and randomly generated. Specifically, each appointment duration is generated as a realization of a discrete uniform random variable in $(0, 1]$. The average and standard deviation of the ratios $\frac{ALG(I)}{OPT(I)}$ are recorded across all 1,000 sequences for the specified experiment setting. Figure 4-14 indicates that the HR algorithm achieves near-optimal performance on average. Moreover, the average ratio $\frac{ALG(I)}{OPT(I)}$ is observed to decrease with larger n. As can be seen, the typical performance of the HR algorithm is within a few percentages of the optimal offline solution.

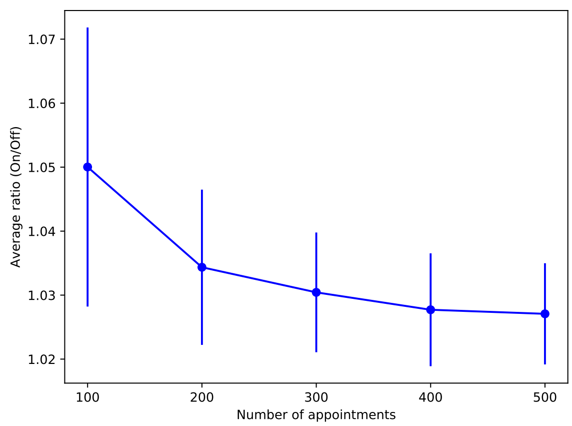

Figure 4-14: Average Competitive Ratio of the HR Algorithm With Associated Standard Deviations.

4.6 Conclusions

This study is the first to formally describe fundamental online scheduling models that capture a broad range of practical applications and specifically the basic setting of scheduling appointments into a capacity limited resource. The study provides strong theoretical contributions that highlight the benefit of flexibility in scheduling appointment interval into physical resources like chairs. In particular, the main theoretical result is that by allowing such flexibility it is possible to design competitive online algorithms with theoretical guarantees that are superior. Specifically, the problem studied in this work could be viewed as a version of the classical bin packing problem with added flexibility. This study describes and analyzes a new algorithm that achieves an asymptotic competitive ratio of 1.5 which is strictly better than the existing lower bound on all randomized algorithms to the online bin packing problem. This highlights the fact that the additional flexibility makes the new model fundamentally different than the bin packing problem.

Finally, the empirical performance of the algorithm seems to be within a few percentages of the offline optimal solution. This implies that the algorithm provides practical solutions at scale. Moreover, the algorithm principles provide important insights that can inform scheduling practices in real life situations.

Chapter 5

Conclusion

This thesis offers actionable insights driven by data and analytics to address key healthcare challenges, including mitigating physicians' professional distress and reducing resource requirements within health systems.

Handling EHR inbox messages is an important driver of work burden for physicians and has been identified as a factor associated with burnout. Chapter 2 first shows that a Latent Dirichlet Allocation model provides a reliable and scalable methodology to analyze the specific content of inbox messages in great level of detail. One of the main conclusion of this study is that content matters as much as quantity, as a consequent part of the time spent by PCPs interacting with other care team members and patients through the inbox is devoted to handling administrative matters. The great number and variety of topics managed by PCPs also highlights the complexity of the nature of their work. Finally, variability across practices suggests that more effort should be devoted to understanding practice level drivers leading to increased administrative burden. This study therefore suggests that inbox workflows redesign could be implemented based on this methodology in order to alleviate PCPs' work burden by better engaging the whole team, leading to improved care efficiency.

Chapter 3 then employs a novel approach to examine team dynamics and their impact on PCP well-being through a quantitative methodology using social network research modeling and available EHR data. Great variability was observed in terms of support team structure and dynamics across PCPs and practices. Factors related to the structure, dynamics and workflows of the PCP's support team were selected as important predictors of different dependent variables, highlighting that PCP's well-being is a function of the overall system surrounding the PCP, rather than features of the physician themselves. Future research should focus on operational system solutions relying on this new methodology.

Chapter 4 finally develops a new modeling framework for the ubiquitous problem of real-time appointment scheduling, called the minimum peak appointment scheduling (MPAS) problem. It captures a broad range of practical applications, and specifically the basic setting of scheduling appointments into a capacity limited resource. This chapter provides strong theoretical contributions that highlight the benefit of flexibility in scheduling appointment interval into physical resources like chairs. Specifically, the MPAS problem can be viewed as a version of the classical bin packing problem with added flexibility. The chapter describes and analyzes a new harmonic rematching algorithm that achieves an asymptotic competitive ratio of 1.5 which is strictly better than the existing lower bound on all randomized algorithms to the online bin packing problem. This highlights the fact that the additional flexibility makes the MPAS problem fundamentally different than the bin packing problem. Finally, the algorithm principles provide important insights that can inform scheduling practices in real life situations.

Appendix A

Work Themes Analysis

A.1 Methodology to Evaluate the Accuracy of the LDA Algorithm

The methodology used to evaluate the accuracy of the predictions made by the LDA model for a given message is presented here. Errors made by the model when predicting a topic are not given the same importance. For example, an error confusing the *Upper-Respiratory* topic with the *Paperwork* topic is considered more significant and is weighted higher than one that confuses the *Test Result Information* topic with the *Blood Test* Topic. Therefore, for each topic, close topics in terms of their main theme are selected among all the topics. For example, the close topics of *Test Result Information* topic will be *Blood Test* topic and *Imaging Exam* topic. The complete methodology used to evaluate the accuracy of the prediction made by the algorithm for a given message is presented below:

1. Initialize accuracy at 0.

2. Initialize the considered topic as the highest-percentage topic of the topic.

3. For the considered topic of the message:

   (a) if the topic is considered present in the message by the physician:

    i. if its associated percentage is considered accurate, the given percentage is added to the accuracy
    ii. if the associated percentage is considered too large, only the deemed accurate percentage is added to the accuracy.
    iii. if the associated percentage is considered too small in comparison to the accurate percentage, only the predicted percentage is added to the accuracy.

  (b) If the topic is not considered present in the message by the physician:
    i. if one of its close topics is considered present in the message, then half of the corresponding predicted percentage is added to the accuracy.
    ii. if none of its close topics are considered present in the message, do nothing.

4. Iterate step 2 for all topics present in the message.

A.2 Topic Labelling

Table A.1: Topic Labelling and Word Distribution

Topic Description	Top Ten Most Frequent Words
Imaging Exam	Scan, ultrasound, breast, cyst, biopsy, cancer, liver, imaging, nodule, kidney
Miscellaneous Chronic Condition	Physical therapy, cardiologist, stress, heart, condition, change, ongoing, diagnosis, facility, therapy
Patient-Initiated Care	Sleep, cold, try, help, counter, using, take, medication, allergy, night
Nutrition	Food, diet, weight, eat, cholesterol, low, high, level, vitamin, supplement, eating

Medical Referral	See, help, could, someone, think, doctor, need, recommend, specialist, refer
Scheduling Appointment	Appointment, see, schedule, visit, office, call, need, time, appt, follow
Upper-Respiratory Symptom	Cough, headache, sinus, symptom, allergy, nose, throat, fever, inhaler, still
Skin Symptom	Skin, red, cream, rash, face, eye, area, picture, dermatologist, itchy
Scheduling Uncompleted Calls	Call, office, back, today, please, phone, need, told, called, said
Ambiguous Feelings	Think, time, feel, see, going, better, feeling, something, night, try
Prescription Administrative Problem	Mg, example, prescription, day, send, medication, tablet, name, pharmacy, supply
Family Referral	Mother, care, mom, patient, family, told, support, husband, friend, year
Joint Procedure	Knee, hip, surgery, car, right, replacement, injury, exercise, walking, accident
Musculoskeletal Symptoms	Hand, neck, right, arm, shoulder, left, pain, side, finger, nerve
Administrative Referral	Referral, discussed, specialist, requesting following, specialist name, need, appointment, fax, request, please
Insurance	Insurance, shot, vaccine, cover, pay, insurance company, new, cost, coverage, cancer

General Paperwork	Non urgent, letter, need, question, send, email, medical, record, work, mail
Medication	Taking, mg, medication, dose, pill, blood pressure, daily, increase, reading, take
Blood Test	Blood, lab, test, order, done, work, level, thyroid, need, go
Refill Coordination	Prescription, pharmacy, refill, need, thanks, script, send, new, request, please
Miscellaneous Identified Symptoms	Stool, diarrhea, symptom, discharge, white, negative, virus, smear, blood, bacterial
Respiratory Symptoms	Antibiotic, ear, chest, infection, lung, pneumonia, ray, prednisone, lesion, patch
Ambiguous Symptoms	Pain, back, symptom, still, leg, feel, non-urgent, muscle, lower, worse
General Check-In	Well, hope, good, great, weekend, best, nice, thanks, see, time
Forms	Forms, need, fill, fax, paperwork, health, complete, school, information, filled
Test Result	Result, test, question, normal, resulted, urine, blood, sample, lab, negative
Feet-Related Problem	Foot, xray, bone, arthritis, toe, fracture, flare, big, needle, big
Miscellaneous Paperwork	Information, patient, behalf, report, note, received, contact, state, dear, regarding
Chronic Cardiovascular and Diabetes Condition	Risk, dl, period, year, unit, age, insulin, treatment, machine, attack

Surgical Procedure	Surgery, procedure, surgeon, orthopedic, endoscopy, month, post, reflux, finish, need

A.3 Sample Messages and Corresponding Topic Mixtures

Sample messages are presented with the topic mixtures and corresponding percentages predicted by the algorithm.

"Dr. X, I want to keep you informed about my current state. Since Friday I have been taking Celexa 30 mg at dinnertime. I seem to be experiencing periods of 'jitteriness' on and off throughout the day. You have described that feeling as a side effect which should 'pass in a few days'. I'm able to manage the feeling so far but wanted you to know that it was occurring. Thank you for your help and support."

Topic mixture: Ambiguous Feelings (60%) + Medication (25%) + none-specific topic (15%)

"Patient states the pain on her right arm is getting worse. She was taking tylenol max mg/day, 6 times a day but she stopped because she thought she was taking too much of it. Denies swelling. Hurts at all times. It also hurts from above the shoulder, on the right side. She can move the right arm, but it is very painful. Pain is also taking away her appetite."

Topic mixture: Ambiguous Pain Symptom (50%) + Musculoskeletal Symptom (27%) + Medication (17%) + none-specific topic (6%)

"I think I'm getting a sinus infection. I have very thick yellow mucus coming from my nose and I have some sinus pressure under my eyes and above. I was wondering if I should schedule an appointment with you in the coming days. Thanks"

Topic mixture: Upper-Respiratory Symptoms (64%) + Scheduling Appointment

(25%) + Miscellaneous Identified Symptoms (6%) + none-specific topic (5%)

"Your labs are back, and for the most part normal. Your urinalysis and culture were negative. Your diabetes screen was negative. Your LDL cholesterol level came back elevated. As a reminder, LDL cholesterol, or "bad" cholesterol is affected by animal fat. That includes red meat and dairy products. Cutting back or eliminating these foods, should help lower your cholesterol, as well as increased exercise. Your Vitamin D level did come back low. Ideally your level should be greater than 20. I would recommend a daily supplement of Vitamin D 1000 U. I would recommend a recheck of you level in 6-12 months. Please let me know if you have any questions or concerns regarding your test result(s)."

Topic mixture: Nutrition (50%) + Test Result (26%) + Blood Test (9%) + none-specific topic (15%)

"Is it possible to send another letter confirming I had my annual exam with the full date of my appointment (MGH Complio rejected the letter you gave me yesterday)? I believe the date of my appointment was December 20, 2017. Thank you!"

Topic mixture: Scheduling Appointment (65%) + General Paperwork (30%) + none-specific topic (5%)

"Patient insurance will not cover methocarbamol (Robaxin) 500 mg TID Insurance will cover 750 mg tab. Please advise on possible prescribing of 750 mg."

Topic mixture: Prescription Administrative Problem (60%) + Insurance (35%) + none-specific topic (5%)

A.4 Topic Description

Table A.2: Description of the Percentage of the Main Topic per Topic

Topic description	Mean of the percentage of the main topic	Standard deviation of the percentage of the main topic
Imaging Exam	0.09	0.13
Miscellaneous Chronic Condition	0.34	0.13
Patient-Initiated Care	0.37	0.13
Nutrition	0.38	0.14
Medical Referral	0.42	0.15
Scheduling Appointment	0.48	0.17
Upper-Respiratory Symptom	0.40	0.13
Skin Symptom	0.37	0.13
Scheduling Uncompleted Calls	0.48	0.17
Ambiguous Feelings	0.40	0.14
Prescription Administrative Problem	0.47	0.17
Family Referral	0.37	0.14
Joint Procedure	0.34	0.12
Musculoskeletal Symptoms	0.37	0.13
Administrative Referral	0.51	0.17
Insurance	0.39	0.14
General Paperwork	0.48	0.15
Medication	0.44	0.15
Blood Test	0.45	0.16
Refill Coordination	0.47	0.16
Miscellaneous Identified Symptoms	0.38	0.13
Respiratory Symptoms	0.45	0.13
Ambiguous Symptoms	0.38	0.13
General Check-In	0.43	0.16
Forms	0.40	0.14
Test Result	0.45	0.17
Feet-Related Problem	0.37	0.13
Miscellaneous Paperwork	0.36	0.12
Chronic Cardiovascular and Diabetes Condition	0.35	0.13
Surgical Procedure	0.36	0.13

Appendix B

Predictive Model for Physician Well-Being

B.1 Methodology to Calculate Network-Based Teamwork Features

- The PCP's *entropy* is calculated through the following formula:

$$entropy = -\sum_{e \in edges: PCP \in e} w_e * \log(w_e)$$

  where w_e is the normalized weight of an edge in the graph. For example, two PCPs connected to 3 staff members with the same global weight of 30 do not have the same entropy: the first PCP with weighted interactions of 28, 1 and 1 has an entropy equal to 0.29, while the second PCP with weighted interactions of 10, 10, and 10 has an entropy of 1.08.

- The *betweenness centrality* of a node v in a weighted graph is calculated as follows:

$$betweenness\ centrality = \sum_{s,t \in V} \frac{\sigma(s,t|v)}{\sigma(s,t)}$$

Where $\sigma(s,t)$ is the number of shortest paths between node s and node t, and where $\sigma(s,t|v)$ is the number of those paths passing through v. Here, the algorithm takes into account the distance between two nodes when calculating the shortest paths. This distance is equal to the inverse of the weight of the edge.

- The *closeness centrality* of a node v in a weighted graph is calculated as follows:

$$closeness\ centrality = \frac{n-1}{\sum_{u \in V, u \neq v} d(u,v)}$$

where $d(u,v)$ is the distance between node u and node v, and n is the number of nodes in the graph. The distance between two nodes is equal to the inverse of the weight of the edge between those two nodes.

- *Turnover* is calculated by comparing the PCP's support team from two adjacent 3-month periods (e.g., March-May 2018 and April-June 2018). The ratio of staff which left the PCP's support team between the two periods is calculated. A large ratio value at a given period indicates that practice turnover impacted the considered PCP. The final turnover feature is created by calculating the third quantile of all the turnover estimates over the different time periods for each PCP. The motivation behind using the third quantile instead of the mean or median is to differentiate PCPs who were especially impacted by turnover compared to an average turnover phenomenon affecting all PCPs.

B.2 Summary Statistics

Table B.1: Summary Statistics for Independent and Dependent Variables

Type of Feature	**Feature**	**Mean ± standard deviation (Median)**	**Range**	**First Quantile**	**Third Quantile**
Structure of support team	Number of nodes	5.63 ± 2.13 (5.10)	2.2 – 12.6	4.1	6.7
	Number of edges	5.82 ± 2.57 (5.35)	1.2 – 16.7	4.0	7.2
	Weight	964.2 ± 646.6 (834.8)	73.3 – 4073.4	483.0	1369.3
	Proportion of nurses	0.44 ± 0.13 (0.45)	0 – 0.79	0.36	0.51
	Proportion of FD	0.12 ± 0.12 (0.08)	0 – 0.45	0	0.21
	Proportion of MAs	0.02 ± 0.06 (0)	0 – 0.33	0	0
	PCP betweenness centrality	0.67 ± 0.32 (0.73)	0 - 1	0.38	0.86
	PCP closeness centrality	159.0 ± 138 (140)	8.9 – 938.9	60.1	201.0
	PCP Entropy	1.18 ± 0.46 (1.15)	0.05 – 2.26	0.87	1.48
	Turnover	0.14 ± 0.10 (0.17)	0 – 0.33	0	0.2

Inbox work theme allocation	Scheduling	0.18 ± 0.03 (0.17)	0.09 – 0.27	0.15	0.19
	Paperwork	0.10 ± 0.02 (0.09)	0.05 – 0.29	0.08	0.11
	Prescription	0.05 ± 0.02 (0.05)	0.02 – 0.10	0.04	0.06
	Administrative Referral	0.02 ± 0.01 (0.02)	0.01 – 0.08	0.01	0.02
	Identified Symptoms	0.07 ± 0.01 (0.07)	0.04 – 0.13	0.06	0.08
	Ambiguous Diagnosis	0.13 ± 0.03 (0.13)	0.07 – 0.22	0.10	0.16
	Condition Management	0.13 ± 0.02 (0.14)	0.08 – 0.19	0.12	0.15
	Clinical Decision-Making Referral	0.08 ± 0.01 (0.08)	0.04 – 0.14	0.07	0.09
	Tests and Exams	0.08 ± 0.02 (0.08)	0.04 – 0.14	0.07	0.09
Inbox Dynamics	Length doctor patient	18.8 ± 8.0 (18)	5 - 47	13.0	22.7
	Length doctor nurse	45.2 ± 33.9 (40)	0 - 219	31	51
	Length nurse doctor	18.1 ± 11.4 (16)	0 - 65	10	25

Dependent Variables	Exhaustion	17.5 ± 7.7 (19)	1 - 30	11	24
	Cynicism	12.1 ± 7.8 (10)	0 - 30	6	18
	Personal Accomplishment	27.6 ± 6.1 (29)	9 - 36	24	32
	Vigor at work	11.4 ± 4.2 (11)	0 - 18	9	15
	Absorption at work	13.4 ± 3.7 (14)	0 - 18	12	16
	Dedication at work	13.5 ± 3.7 (14.5)	0 - 18	11	16
	Professional fulfillment	14.5 ± 5.3 (14.5)	0 - 24	11	19
	Perceived appreciation	13.1 ± 4.9 (13)	3 - 24	9.7	17
	Peer support	9.7 ± 4.0 (10)	0 - 16	7	12

Table B.2: Comparison of the Feature Characteristics between Survey Respondents and Non-Respondents

Type of feature	**Name of feature**	**PCPs who answered the survey (Mean ± Standard Deviation)**	**PCPs who did not answer the survey (Mean ± Standard Deviation)**	**Pvalue**

Structure of support team	Number of nodes	5.63 ±2.13	5.84±1.66	0.58
	Number of edges	5.82 ±2.57	6.24±2.37	0.41
	Weight	964.2 ±646.6	820.6±659	0.30
	Proportion of nurses	0.44±0.13	0.42±0.14	0.46
	Proportion of FD	0.12±0.12	0.12±0.11	0.87
	Proportion of MAs	0.02±0.06	0.05±0.10	0.10
	PCP Entropy	1.18±0.46	1.22±0.30	0.58
	PCP betweenness centrality	0.67±0.32	0.65±0.25	0.43
	PCP closeness centrality	159.0±138	114.1±112.4	0.08
	Turnover	0.14±0.10	0.18±0.03	0.56
Inbasket work theme allocation	Scheduling	0.18±0.03	0.18±0.03	0.94
	Paperwork	0.10±0.02	0.09±0.02	0.90
	Prescription	0.05±0.02	0.05±0.02	0.31
	Administrative Referral	0.02±0.01	0.02±0.02	0.92
	Identified Symptoms	0.07±0.01	0.06±0.02	0.04

	Ambiguous Diagnosis	0.13±0.03	0.14±0.03	0.12
	Condition Management	0.13±0.02	0.12±0.02	0.06
	Clinical Decision-Making Referral	0.08±0.01	0.09±0.02	0.04
	Test and Exam	0.08±0.02	0.09±0.03	0.60
Inbasket dynamics	Length doctor patient	18.8±8.0	19.9±5.5	0.40
	Length doctor nurse	45.2±33.9	57.3±38.5	0.14
	Length nurse doctor	18.1±11.4	23.0±14.5	0.11

B.3 Predictive Results

Table B.3: Features Distribution Across Practices

Practice	Fraction of nurses	Fraction of MAs	Fraction of FD	Entropy	Betweenness	Turnover
A	0.12 (0.07)	0.01 (0.02)	0.43 (0.03)	0.97 (0.08)	0.30 (0.17)	0.09 (0.11)
B	0.48 (0.03)	0.01 (0.03)	0.06 (0.05)	1.06 (0.23)	0.15 (0.14)	0.06 (0.10)
C	0.15 (0.11)	0.29 (0.03)	0.00 (0.00)	0.95 (0.07)	0.52 (0.34)	0.16 (0.11)
D	0.39 (0.06)	0.02 (0.03)	0.38 (0.04)	1.89 (0.14)	0.67 (0.10)	0.19 (0.06)
E	0.37 (0.07)	0.00 (0.01)	0.27 (0.04)	1.38 (0.16)	0.67 (0.18)	0.13 (0.09)
F	0.64 (0.09)	0.00 (0.00)	0.07 (0.10)	1.42 (0.17)	0.85 (0.07)	0.19 (0.09)
G	0.59 (0.08)	0.05 (0.06)	0.12 (0.09)	1.56 (0.23)	0.93 (0.06)	0.20 (0.04)
H	0.43 (0.13)	0.00 (0.00)	0.21 (0.06)	1.83 (0.24)	0.97 (0.04)	0.16 (0.06)
I	0.48 (0.03)	0.00 (0.00)	0.00 (0.00)	0.80 (0.17)	0.01 (0.02)	0.09 (0.12)
J	0.58 (0.12)	0.00 (0.00)	0.11 (0.10)	1.62 (0.30)	0.88 (0.07)	0.12 (0.06)
K	0.45 (0.09)	0.00 (0.00)	0.02 (0.05)	0.76 (0.28)	0.47 (0.30)	0.08 (0.13)
L	0.44 (0.05)	0.00 (0.00)	0.04 (0.03)	0.68 (0.34)	0.52 (0.26)	0.17 (0.13)
M	0.34 (0.03)	0.24 (0.03)	0.05 (0.05)	1.35 (0.10)	0.54 (0.23)	0.09 (0.10)
N	0.49 (0.06)	0.04 (0.07)	0.09 (0.03)	1.25 (0.13)	0.32 (0.21)	0.18 (0.08)
O	0.46 (0.09)	0.01 (0.02)	0.15 (0.02)	1.30 (0.25)	0.80 (0.11)	0.20 (0.05)
P	0.40 (0.03)	0.00 (0.00)	0.25 (0.07)	1.97 (0.20)	0.85 (0.04)	0.17 (0.08)
Q	0.39 (0.06)	0.00 (0.00)	0.20 (0.05)	1.16 (0.20)	0.82 (0.12)	0.17 (0.08)

Table B.4: Regularized Logistic Regression Prediction Performance for Different Dependent Variables

Dependent variable	**Average Out-of-sample AUC**	**SD (Out-of-sample AUC)**
Cynicism	0.576	0.089
Personal Accomplishment	0.689	0.088
Absorption	0.654	0.089
Dedication	0.655	0.088
Peer Support	0.648	0.083

Table B.5: Selected Feature Coefficients in the Regularized Logistic Regression Model for Different Dependent Variables

Features	**Cynicism**	**Personal Accomplishment**	**Absorption**	**Dedication**	**Peer support**
Gender					-0.01
Graduation Year	-0.30	-0.31			
After-Hour	-0.23		-0.57		
Risk Score	0.31		0.51		
Scheduling		0.32		0.41	
Paperwork					0.01
Ambiguous Diagnosis					-0.01
Fraction MA		-0.41			-0.008
Fraction FD			0.39		
Fraction nurse		-0.49		-0.33	
Length nurse doctor		-0.42	-0.6	-0.51	

Table B.6: Sensitivity to the Threshold Used to Define Exhaustion Binary Outcome

	Threshold 23, 50%	**Threshold 24, 43%**	**Threshold 25, 40%**
ROC score	0.685 (0.087)	0.701	0.681 (0.086)
FTE	0.84	0.91	0.78
Gender			
Graduation year			
After-hour	0.42	0.50	0.42
Risk score			
Scheduling			
Paperwork			
Prescription			
Administrative Referral			
Identified Symptoms			
Ambiguous Diagnosis			
Care Management			
Test			
Medical Referral			
Entropy	-0.54	-0.75	-0.91
Ratio patient doctor			
Fraction MA			
Fraction FD			
Fraction nurse			
Weighted betweenness	0.50	0.53	0.64
Weighted closeness			
Turnover			
Length doctor patient	0.68	0.67	0.24
Length doctor nurse			
Length nurse doctor			

Table B.7: Sensitivity to the Threshold Used to Define Vigor Binary Outcome

	Threshold 15, 55%	**Threshold 14, 46%**
ROC score	0.666 (0.085)	0.692 (0.085)
FTE		
Gender		
Graduation year		
After-hour		
Risk score		
Scheduling		
Paperwork		
Prescription		
Administrative Referral		
Identified Symptoms		
Ambiguous Diagnosis		
Care Management		
Test		
Medical Referral		
Entropy		
Ratio patient doctor		
Fraction MA	-0.11	-0.55
Fraction FD		
Fraction nurse		
Weighted betweenness	0.13	0.47
Weighted closeness		
Turnover	0.07	0.34
Length doctor patient	0.06	0.33
Length doctor nurse		
Length nurse doctor	-0.07	-0.47

Table B.8: Sensitivity to the Threshold Used to Define Professional Fulfillment Binary Outcome

	Threshold 21, 50%	**Threshold 20, 44%**
ROC score	0.667 (0.082)	0.630 (0.090)
FTE		
Gender		
Graduation year		
After-hour		
Risk score		
Scheduling	0.43	0.53
Paperwork		
Prescription		
Administrative Referral		
Identified Symptoms		
Ambiguous Diagnosis		
Care Management		
Test		
Medical Referral		
Entropy		
Ratio patient doctor		
Fraction MA		
Fraction FD		
Fraction nurse		
Weighted betweenness		
Weighted closeness		
Turnover	0.10	0.08
Length doctor patient		
Length doctor nurse		
Length nurse doctor	-0.40	-0.44

Appendix C

Proof of Proposition 2

In order to show a lower bound on the absolute competitive ratio of all online algorithms, the following lemma considers the manner by which the first appointment is scheduled.

Lemma 11. *Any online algorithm that does not schedule the first appointment at either time 0 or time $T - \ell_1$ (i.e., $s_1 \neq 0, T - \ell_1$) has an absolute competitive ratio ≥ 2.*

Proof. Let's write l_1 the length of the first appointment. Let r_1 and r_2 the remaining available time before and after A_1, respectively. Thus, $r_1 = s_1 > 0$ and $r_2 = T - s_1 - \ell_1 > 0$. Let $l_2 = \min(r_1, r_2) + \epsilon$ and $l_3 = \max(r_1, r_2) - \epsilon$. In this configuration, it is impossible to fit both A_2 and A_3 in the same bin as A_1 even with the added flexibility. However, it is easy to check that $l_1 + l_2 + l_3 = 1$, and so OPT could fit all the jobs in the same bin. Therefore, the absolute competitive ratio of any online algorithm that does not schedule the first appointment at either time 0 or time $T - \ell_1$ (i.e., $s_1 \neq 0, T - \ell_1$) must be bigger than 2.

□

The proof of the main result follows.

Proof. The proposition is proved by construction. In particular, a dynamic problem instance is developed that observes the incremental scheduling decisions made by any online algorithm (randomized or not) and adjusts the subsequent appointment requests in an adversarial manner that causes the online algorithm to use unnecessary chairs once future appointment requests are realized. For each step, all possible changes in the appointment-to-chair assignment allowed by the additional flexibility of the MPAS problem are considered.

Without loss of generality, discretized time slots will be considered and time horizon will be equal to $T = 20$ instead of $T = 1$. It will also be considered that one appointment can be processed by multiple chairs, which is a more flexible setting than the one developed in Section 4.3. Therefore, the result on the lower bound obtained here will still be true in the more constrained setting considered in Section 4.3.

Let $\ell_1 = 10$, and $\ell_2 = 9$. By Lemma 11, one only needs to consider the cases where $s_1 = 0$ or $s_1 = 10$ because the online algorithm would otherwise immediately have a competitive ratio lower bounded by 2. Assume that $s_1 = 0$ (by symmetry, the case where $s_1 = 10$ follows analogously). Now, consider the two cases that do not immediately yield a lower bound of 2 on the absolute competitive ratio:

Case 1: $s_2 = 10$. Introduce three additional appointments: $\ell_3 = 2$, $\ell_4 = 10$, and $\ell_5 = 9$. Then, it is not difficult to see in Figure C-1 that no matter what s_3, s_4, and s_5 are chosen, the resulting peak utilization will be at least 3.

A1 (s=0, L=10)										**A2 (s=10, L=9)**									
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20

Figure C-1: Proposition 2: Case 1

Case 2: $s_2 = 11$. Introduce a third appointment $\ell_3 = 2$ and condition on how the online algorithm schedules this appointment.

Case 2.1: $s_3 = 0$ or 18. Introduce a fourth appointment $\ell_4 = 2$ and condition on how the online algorithm schedules this appointment.

Case 2.1.1: A_4 is scheduled back-to-back (either immediately before or after) with A_3. Thus, if $s_3 = 0$ then $s_4 = 2$, else if $s_3 = 18$ then $s_4 = 16$. Then introduce a fifth and sixth appointment: $\ell_5 = 8, \ell_6 = 9$. These last two appointments clearly cannot be added to the existing schedules shown in Figures C-2 and C-3 without increasing the peak utilization to 3. However, an offline algorithm could have used only 2 chairs by placing A_1, A_3, and A_5 in one chair and A_2, A_4, and A_6 in the other chair.

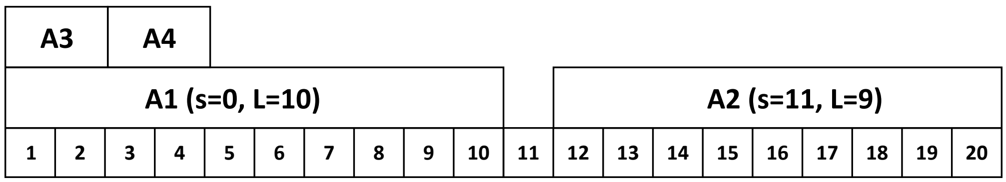

Figure C-2: Proposition 2: Case 2.1.1; $s_3 = 0$

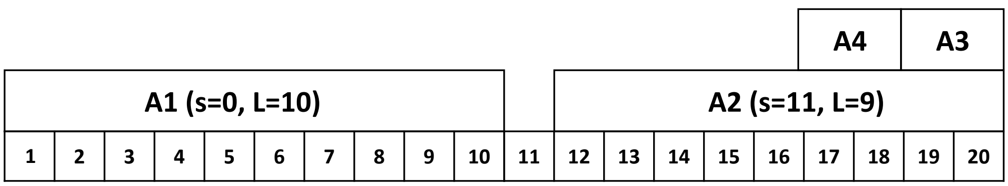

Figure C-3: Proposition 2: Case 2.1.1; $s_3 = 18$

Case 2.1.2: $s_3 = 0$ and $s_4 = 9$. Then introduce a fifth and sixth appointment: $\ell_5 = 8, \ell_6 = 9$. These last two appointments clearly cannot be added to the existing schedule shown in Figure C-4 without increasing the peak utilization to 3. However, an offline algorithm could have used only 2 chairs by placing A_1, A_3, and A_5 in one chair and A_2, A_4, and A_6 in the other chair.

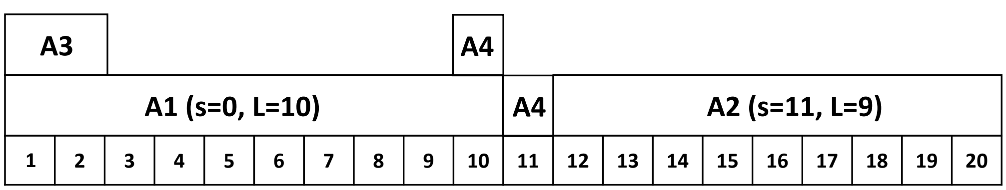

Figure C-4: Proposition 2: Case 2.1.2

Case 2.1.3: All other instances where A_4 is not scheduled back-to-back with A_3 and $(s_3, s_4) \neq (0, 9)$. Then, introduce a fifth and sixth appointment: $\ell_5 = 10, \ell_6 = 7$. These last two appointments clearly cannot be added to the existing schedules illustrated in Figures C-5, C-6, C-7, and C-8 without increasing the peak utilization to 3. However, an offline algorithm could have used only 2 chairs by placing A_1 and A_5 in one chair and A_2, A_3, A_4, and A_6 in the other chair.

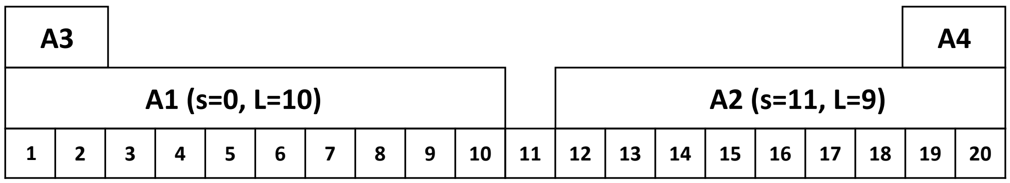

Figure C-5: Proposition 2: Case 2.1.3; Scenario a

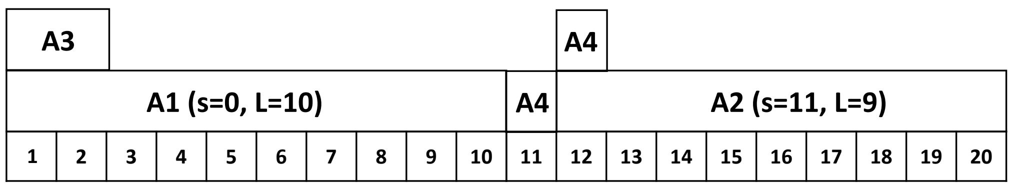

Figure C-6: Proposition 2: Case 2.1.3; Scenario b

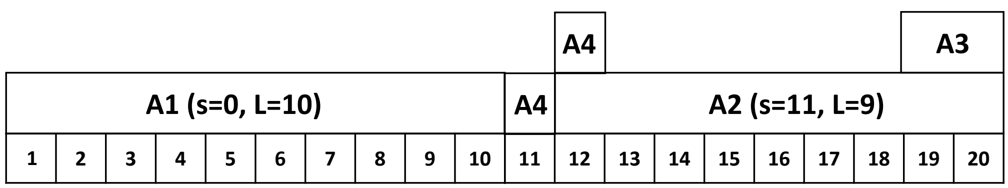

Figure C-7: Proposition 2: Case 2.1.3; Scenario c

Case 2.2: $s_3 = 9$ or 10. Introduce a fourth appointment $\ell_4 = 2$. Regardless of where the fourth appointment is scheduled, there are at least two strings of consecutive unoccupied time slots on the second row of the histogram as shown in Figures C-9 and C-10. Denote the length of the longest such string as e. If $e < 10$, then introduce a fifth appointment $\ell_5 = 10$. Else, introduce a fifth and sixth appointment $\ell_5 = 8$ and $\ell_6 = 9$. In either scenario, the online algorithm will be forced to use 3 chairs while the offline algorithm only uses 2.

Case 2.3: $s_3 \neq 0, 9, 10, 18$. Then, introduce a fourth and fifth appointment $\ell_5 = 10, \ell_6 = 9$. The online algorithm is then forced to use 3 chairs while the offline algorithm only uses 2. □

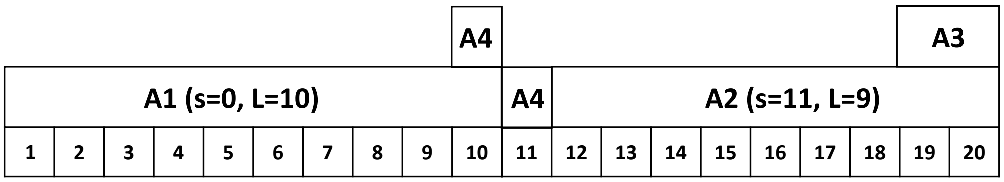

Figure C-8: Proposition 2: Case 2.1.3; Scenario d

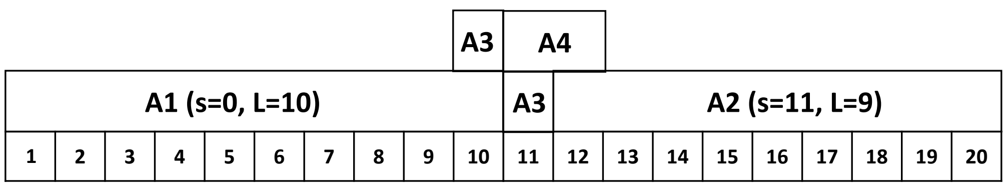

Figure C-9: Proposition 2: Case 2.2; $s_3 = 9$

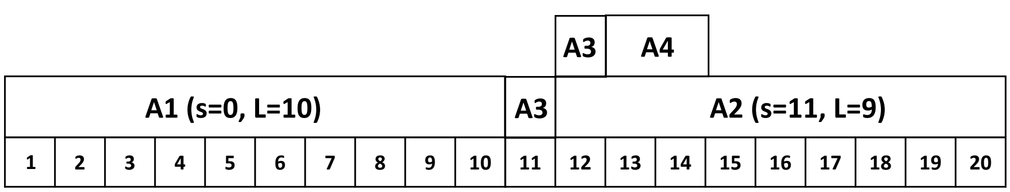

Figure C-10: Proposition 2: Case 2.2; $s_3 = 10$

Appendix D

Types of One-Sided Bins and Complete Sets of Matched One-Sided Bins

Let's define Type 1 and Type 2 one-sided bins for each job category.

- Job category $(\frac{1}{4}, \frac{1}{3}]$

  Type 1 bins are made of 3 jobs. Type 2 thin one-sided bins are made of 1 job, and Type 2 fat one-sided bins are made of 1 job. A complete set of matched one-sided bins is made of four Type 1 bins, and two Type 2 bins.

- Job category $(\frac{1}{5}, \frac{1}{4}]$

  Type 1 bins are made of 4 jobs. Type 2 thin one-sided bins are made of 1 job, and Type 2 fat one-sided bins are made of 2 jobs. A complete set of matched one-sided bins is made of one Type 1 bin, and two Type 2 bins.

- Job category $(\frac{1}{6}, \frac{1}{5}]$

  Type 1 bins are made of 5 jobs. Type 2 thin one-sided bins are made of 1 job, and Type 2 fat one-sided bins are made of 2 jobs. A complete set of matched one-sided bins is made of two Type 1 bins, and two Type 2 bins.

- Job category $(\frac{1}{7}, \frac{1}{6}]$

  No Type 1 one-sided bins are used for this job category. Type 2 thin one-sided bins are made of 2 jobs, and Type 2 fat one-sided bins are made of 3 job. A complete set of matched one-sided bins is made of two Type 2 bins.

- Job category $(\frac{1}{8}, \frac{1}{7}]$

  Type 1 bins are made of 7 jobs. Type 2 thin one-sided bins are made of 2 jobs, and Type 2 fat one-sided bins are made of 3 jobs. A complete set of matched one-sided bins is made of two Type 1 bins, and ten Type 2 bins.

- Job category $(\frac{1}{9}, \frac{1}{8}]$

  No Type 1 one-sided bins are used for this job category. Type 2 thin one-sided bins are made of 2 jobs, and Type 2 fat one-sided bins are made of 4 jobs. A complete set of matched one-sided bins is made of two Type 2 bins.

- Job category $(\frac{1}{10}, \frac{1}{11}]$

  No Type 1 one-sided bins are used for this job category. Type 2 thin one-sided bins are made of 2 jobs, and Type 2 fat one-sided bins are made of 3 jobs. A complete set of matched one-sided bins is made of two Type 2 bins.

- Job category $(\frac{1}{11}, \frac{1}{10}]$

  No Type 1 one-sided bins are used for this job category. Type 2 thin one-sided bins are made of 3 jobs, and Type 2 fat one-sided bins are made of 4 jobs. A complete set of matched one-sided bins is made of two Type 2 bins.

- Job category $(\frac{1}{12}, \frac{1}{11}]$

  No Type 1 one-sided bins are used for this job category. Type 2 thin one-sided bins are made of 3 jobs, and Type 2 fat one-sided bins are made of 5 jobs. A complete set of matched one-sided bins is made of two Type 2 bins.

- Job category $[0, \frac{1}{12}]$

For this category, a fixed number of jobs cannot be specified per type of bin, as jobs can be infinitely small. Therefore, a Type 2 thin one-sided bin is filled with jobs as long as the level is strictly smaller than $\frac{1}{4}$, and a Type 2 fat one-sided bin is filled with jobs as long as the level is strictly smaller than $\frac{5}{12}$. A complete set of matched one-sided bins is made of two Type 2 bins.

Bibliography

[1] Julia Adler-Milstein, Wendi Zhao, Rachel Willard-Grace, Margae Knox, and Kevin Grumbach. Electronic health records and burnout: Time spent on the electronic health record after hours and message volume associated with exhaustion but not with cynicism among primary care clinicians. *Journal of the American Medical Informatics Association: JAMIA*, 27(4):531–538, April 2020.

[2] Z Ahmed, TY ElMekkawy, and S Bates. Developing an efficient scheduling template of a chemotherapy treatment unit: A case study. *The Australasian medical journal*, 4(10):575, 2011.

[3] Justin Altschuler, David Margolius, Thomas Bodenheimer, and Kevin Grumbach. Estimating a reasonable patient panel size for primary care physicians with team-based task delegation. *The Annals of Family Medicine*, 10(5):396–400, 2012.

[4] Lawrence J Appel, Jeanne M Clark, Hsin-Chieh Yeh, Nae-Yuh Wang, Janelle W Coughlin, Gail Daumit, Edgar R Miller III, Arlene Dalcin, Gerald J Jerome, Steven Geller, et al. Comparative effectiveness of weight-loss interventions in clinical practice. *New England Journal of Medicine*, 365(21):1959–1968, 2011.

[5] Brian G. Arndt, John W. Beasley, Michelle D. Watkinson, Jonathan L. Temte, Wen-Jan Tuan, Christine A. Sinsky, and Valerie J. Gilchrist. Tethered to the EHR: Primary Care Physician Workload Assessment Using EHR Event Log Data and Time-Motion Observations. *The Annals of Family Medicine*, 15(5):419–426, September 2017. Publisher: The Annals of Family Medicine Section: Original Research.

[6] Association of American Medical Colleges. The Complexities of Physician Supply and Demand: Projections from 2018 to 2033: Final Report. https://www.aamc.org/system/files/2020-06/stratcomm-aamc-physician-workforce-projections-june-2020.pdf. Accessed March 23, 2021.

[7] William Atkinson, Anita Misra-Hebert, and James K Stoller. The impact on revenue of physician turnover: an assessment model and experience in a large healthcare center. *The Journal of medical practice management: MPM*, 21(6):351–355, 2006.

[8] Stewart Babbott, Linda Baier Manwell, Roger Brown, Enid Montague, Eric Williams, Mark Schwartz, Erik Hess, and Mark Linzer. Electronic medical records and physician stress in primary care: results from the MEMO Study. *Journal of the American Medical Informatics Association: JAMIA*, 21(e1):e100–106, February 2014.

[9] Arnold B. Bakker, Evangelia Demerouti, and Wilmar B. Schaufeli. Validation of the Maslach Burnout Inventory - General Survey: An Internet Study. *Anxiety, Stress, & Coping*, 15(3):245–260, January 2002. Publisher: Routledge _eprint: https://doi.org/10.1080/10615800210000020716.

[10] Charles M Balch, Michael R Oreskovich, Lotte N Dyrbye, Joseph M Colaiano, Daniel V Satele, Jeff A Sloan, and Tait D Shanafelt. Personal consequences of malpractice lawsuits on american surgeons. *Journal of the American College of Surgeons*, 213(5):657–667, 2011.

[11] J. Balogh, J. Békési, G. Dosa, J. Sgall, and R. van Stee. The optimal absolute ratio for online bin packing. *Journal of Computer and System Sciences*, 2019.

[12] János Balogh, József Békési, György Dósa, Leah Epstein, and Asaf Levin. A new and improved algorithm for online bin packing. *arXiv preprint arXiv:1707.01728*, 2017.

[13] János Balogh, József Békési, György Dósa, Leah Epstein, and Asaf Levin. A new lower bound for classic online bin packing. In *International Workshop on Approximation and Online Algorithms*, pages 18–28. Springer, 2019.

[14] János Balogh, József Békési, and Gábor Galambos. New lower bounds for certain classes of bin packing algorithms. *Theoretical Computer Science*, 440:1–13, 2012.

[15] David M. Blei, Andrew Y. Ng, and Michael I. Jordan. Latent dirichlet allocation. *the Journal of machine Learning research*, 3:993–1022, 2003. Publisher: JMLR. org.

[16] Thomas Bodenheimer, Ellen Chen, and Heather D Bennett. Confronting the growing burden of chronic disease: can the us health care workforce do the job? *Health affairs*, 28(1):64–74, 2009.

[17] Sue Bowman. Impact of electronic health record systems on information integrity: quality and safety implications. *Perspectives in health information management*, 10(Fall), 2013.

[18] Ulrik Brandes. A faster algorithm for betweenness centrality. *Journal of mathematical sociology*, 25(2):163–177, 2001.

[19] Sharon Bell Buchbinder, Modena Wilson, Clifford F Melick, Neil R Powe, et al. Estimates of costs of primary care physician turnover. *Am J Manag Care*, 5(11):1431–1438, 1999.

[20] Ruth A Bush, Cynthia D Connelly, Alexa Pérez, Neilson Chiang, Cynthia Kuelbs, and George J Chiang. Physician perception of the role of the patient portal in pediatric health. *The Journal of ambulatory care management*, 40(3):238, 2017.

[21] Centers for Medicare & Medicaid Services. National health expenditure accounts (NHEA). Technical report, U.S. Federal Government, 2018.

[22] Barun Chandra. Does randomization help in online bin packing ? *Information processing letters*, 43:15–19, 1992.

[23] Jonathan Chang, Jordan Boyd-Graber, Chong Wang, Sean Gerrish, and David M. Blei. Reading tea leaves: How humans interpret topic models. In *Neural information processing systems*, volume 22, pages 288–296. Citeseer, 2009.

[24] Kuan-Yu Chen, Che-Ming Yang, Che-Hui Lien, Hung-Yi Chiou, Mau-Roung Lin, Hui-Ru Chang, and Wen-Ta Chiu. Burnout, job satisfaction, and medical malpractice among physicians. *International Journal of Medical Sciences*, 10(11):1471–1478, 2013.

[25] Melinda A Chen, James P Hollenberg, Walid Michelen, Janey C Peterson, and Lawrence P Casalino. Patient care outside of office visits: a primary care physician time study. *Journal of general internal medicine*, 26(1):58–63, 2011.

[26] Edward G Coffman and János Csirik. Performance guarantees for one-dimensional bin packing. In *Handbook of approximation algorithms and metaheuristics*, pages 32–1. CRC Press, 2007.

[27] Edward G Coffman, János Csirik, Gábor Galambos, Silvano Martello, and Daniele Vigo. Bin packing approximation algorithms: survey and classification. In *Handbook of combinatorial optimization*, pages 455–531. Springer New York, 2013.

[28] Edward G Coffman, Gabor Galambos, Silvano Martello, and Daniele Vigo. Bin packing approximation algorithms: Combinatorial analysis. In *Handbook of combinatorial optimization*, pages 151–207. Springer, 1999.

[29] János Csirik and Gerhard J Woeginger. On-line packing and covering problems. In *Online Algorithms*, pages 147–177. Springer, 1998.

[30] Matthias Dehmer and Abbe Mowshowitz. A history of graph entropy measures. *Information Sciences*, 181(1):57–78, 2011.

[31] Carolyn S Dewa, Desmond Loong, Sarah Bonato, Nguyen Xuan Thanh, and Philip Jacobs. How does burnout affect physician productivity? a systematic literature review. *BMC health services research*, 14(1):1–10, 2014.

[32] M Robin DiMatteo, Cathy Donald Sherbourne, Ron D Hays, Lynn Ordway, Richard L Kravitz, Elizabeth A McGlynn, Sherrie Kaplan, and William H Rogers. Physicians' characteristics influence patients' adherence to medical treatment: results from the medical outcomes study. *Health psychology*, 12(2):93, 1993.

[33] Patrick M Dunn, Bengt B Arnetz, John F Christensen, and Louis Homer. Meeting the imperative to improve physician well-being: assessment of an innovative program. *Journal of general internal medicine*, 22(11):1544–1552, 2007.

[34] Liselotte N Dyrbye, Matthew R Thomas, F Stanford Massie, David V Power, Anne Eacker, William Harper, Steven Durning, Christine Moutier, Daniel W Szydlo, Paul J Novotny, et al. Burnout and suicidal ideation among us medical students. *Annals of internal medicine*, 149(5):334–341, 2008.

[35] Shari M. Erickson, Brooke Rockwern, Michelle Koltov, Robert M. McLean, and Medical Practice and Quality Committee of the American College of Physicians. Putting Patients First by Reducing Administrative Tasks in Health Care: A Position Paper of the American College of Physicians. *Annals of Internal Medicine*, 166(9):659–661, May 2017.

[36] Donald K Freeborn. Satisfaction, commitment, and psychological well-being among hmo physicians. *Western Journal of Medicine*, 174(1):13, 2001.

[37] Ann Fridner, Karen Belkić, Daria Minucci, Luigi Pavan, Massimo Marini, Birgit Pingel, Giovanni Putoto, Pierluigi Simonato, Lise T Løvseth, and Karin Schenck-Gustafsson. Work environment and recent suicidal thoughts among male university hospital physicians in sweden and italy: the health and organization among university hospital physicians in europe (houpe) study. *Gender medicine*, 8(4):269–279, 2011.

[38] Jerome Friedman, Trevor Hastie, and Rob Tibshirani. Regularization paths for generalized linear models via coordinate descent. *Journal of statistical software*, 33(1):1, 2010.

[39] Jerome Friedman, Trevor Hastie, Robert Tibshirani, et al. *The elements of statistical learning*, volume 1. Springer series in statistics New York, 2001.

[40] Gabor Galambos and JBG Frenk. A simple proof of liang's lower bound for on-line bin packing and the extension to the parametric case. *Discrete Applied Mathematics*, 41(2):173–178, 1993.

[41] Rebekah L. Gardner, Emily Cooper, Jacqueline Haskell, Daniel A. Harris, Sara Poplau, Philip J. Kroth, and Mark Linzer. Physician stress and burnout: the impact of health information technology. *Journal of the American Medical Informatics Association: JAMIA*, 26(2):106–114, February 2019.

[42] Valerie Gilchrist, Gary McCord, Susan Labuda Schrop, Bridget D. King, Kenelm F. McCormick, Allison M. Oprandi, Brian A. Selius, Michael Cowher, Rishi Maheshwary, Falguni Patel, Ami Shah, Bonny Tsai, and Mia Zaharna. Physician Activities During Time Out of the Examination Room. *The Annals of Family Medicine*, 3(6):494–499, November 2005. Publisher: The Annals of Family Medicine Section: Original Research.

[43] Paul C Gilmore and Alan J Hoffman. A characterization of comparability graphs and of interval graphs. *Canadian Journal of Mathematics*, 16:539–548, 1964.

[44] Martin Charles Golumbic. *Algorithmic graph theory and perfect graphs*. Elsevier, 2004.

[45] R Grol, H Mokkink, A Smits, J Van Eijk, M Beek, P Mesker, and J Mesker-Niesten. Work satisfaction of general practitioners and the quality of patient care. *Family practice*, 2(3):128–135, 1985.

[46] Kevin Grumbach and Thomas Bodenheimer. A primary care home for Americans: putting the house in order. *JAMA*, 288(7):889–893, August 2002.

[47] Isabelle Guyon, Jason Weston, Stephen Barnhill, and Vladimir Vapnik. Gene selection for cancer classification using support vector machines. *Machine learning*, 46(1):389–422, 2002.

[48] David Hall, Daniel Jurafsky, and Christopher D. Manning. Studying the History of Ideas Using Topic Models. In *Proceedings of the 2008 Conference on Empirical Methods in Natural Language Processing*, pages 363–371, Honolulu, Hawaii, October 2008. Association for Computational Linguistics.

[49] Caroline Haythornthwaite, Barry Wellman, and Marilyn Mantei. Work relationships and media use: A social network analysis. *Group Decision and Negotiation*, 4(3):193–211, 1995.

[50] Itai Himelboim, Marc A Smith, Lee Rainie, Ben Shneiderman, and Camila Espina. Classifying twitter topic-networks using social network analysis. *Social media+ society*, 3(1):2056305117691545, 2017.

[51] Jenna Howard, Elizabeth C. Clark, Asia Friedman, Jesse C. Crosson, Maria Pellerano, Benjamin F. Crabtree, Ben-Tzion Karsh, Carlos R. Jaen, Douglas S. Bell, and Deborah J. Cohen. Electronic health record impact on work burden in small, unaffiliated, community-based primary care practices. *Journal of General Internal Medicine*, 28(1):107–113, January 2013.

[52] Kirby Hsu and Victor Marshall. Prevalence of depression and distress in a large sample of. *Am J Psychiatry*, 144(12):1561–6, 1987.

[53] Michael Hu. *Leveraging data analytics to improve outpatient healthcare operations*. PhD thesis, Massachusetts Institute of Technology, 2020.

[54] David S Johnson. *Near-optimal bin packing algorithms.* PhD thesis, Massachusetts Institute of Technology, 1973.

[55] David S Johnson. Fast algorithms for bin packing. *Journal of Computer and System Sciences*, 8(3):272–314, 1974.

[56] David S. Johnson, Alan Demers, Jeffrey D. Ullman, Michael R Garey, and Ronald L. Graham. Worst-case performance bounds for simple one-dimensional packing algorithms. *SIAM Journal on computing*, 3(4):299–325, 1974.

[57] John W Jones, Bruce N Barge, Brian D Steffy, Lisa M Fay, Lisa K Kunz, and Lisa J Wuebker. Stress and medical malpractice: organizational risk assessment and intervention. *Journal of Applied Psychology*, 73(4):727, 1988.

[58] Antoon WJ Kolen, Jan Karel Lenstra, Christos H Papadimitriou, and Frits CR Spieksma. Interval scheduling: A survey. *Naval Research Logistics (NRL)*, 54(5):530–543, 2007.

[59] T. R. Konrad, E. S. Williams, M. Linzer, J. McMurray, D. E. Pathman, M. Gerrity, M. D. Schwartz, W. E. Scheckler, J. Van Kirk, E. Rhodes, and J. Douglas. Measuring physician job satisfaction in a changing workplace and a challenging environment. SGIM Career Satisfaction Study Group. Society of General Internal Medicine. *Medical Care*, 37(11):1174–1182, November 1999.

[60] Dana Larsen, Jacqueline T Chu, Liyang Yu, Yuchiao Chang, Karen Donelan, and Kerri Palamara. Correlating burnout and well-being in a multisite study of internal medicine residents and faculty. *Journal of General Internal Medicine*, pages 1–5, 2021.

[61] Chan C Lee and Der-Tsai Lee. A simple on-line bin-packing algorithm. *Journal of the ACM (JACM)*, 32(3):562–572, 1985.

[62] Bohui Liang, Ayten Turkcan, Mehmet Erkan Ceyhan, and Keith Stuart. Improvement of chemotherapy patient flow and scheduling in an outpatient oncology clinic. *International Journal of Production Research*, 53(24):7177–7190, 2015.

[63] Frank M Liang. A lower bound for online bin packing. *Information processing letters*, 10(2), 1980.

[64] EG Co man Jr, MR Garey, and DS Johnson. Approximation algorithms for bin packing: A survey. *Approximation algorithms for NP-hard problems*, pages 46–93, 1996.

[65] Christina Maslach, Susan E. Jackson, Michael P. Leiter, Wilmar B. Schaufeli, and Richard L. Schwab. *Maslach burnout inventory*, volume 21. Consulting psychologists press Palo Alto, CA, 1986.

[66] R Scott McCain, Nicola McKinley, Martin Dempster, W Jeffrey Campbell, and Stephen J Kirk. A study of the relationship between resilience, burnout and coping strategies in doctors. *Postgraduate medical journal*, 94(1107):43–47, 2018.

[67] Clement J McDonald, Fiona M Callaghan, Arlene Weissman, Rebecca M Goodwin, Mallika Mundkur, and Thomson Kuhn. Use of internist's free time by ambulatory care electronic medical record systems. *JAMA internal medicine*, 174(11):1860–1863, 2014.

[68] Clement J. McDonald and Michael H. McDonald. Electronic medical records and preserving primary care physicians' time: comment on "electronic health record-based messages to primary care providers". *Archives of Internal Medicine*, 172(3):285–287, February 2012.

[69] Laurence F McMahon, Kim Rize, NiJuanna Irby-Johnson, and Vineet Chopra. Designed to fail? the future of primary care. *Journal of General Internal Medicine*, 36(2):515–517, 2021.

[70] Arabella Melville. Job satisfaction in general practice implications for prescribing. *Social Science & Medicine. Part A: Medical Psychology & Medical Sociology*, 14(6):495–499, 1980.

[71] Nikitha K Menon, Tait D Shanafelt, Christine A Sinsky, Mark Linzer, Lindsey Carlasare, Keri JS Brady, Martin J Stillman, and Mickey T Trockel. Association of physician burnout with suicidal ideation and medical errors. *JAMA network open*, 3(12):e2028780–e2028780, 2020.

[72] David P Miller Jr, Celine Latulipe, Kathryn A Melius, Sara A Quandt, and Thomas A Arcury. Primary care providers' views of patient portals: interview study of perceived benefits and consequences. *Journal of medical Internet research*, 18(1):e4953, 2016.

[73] Anita D Misra-Hebert, Robert Kay, and James K Stoller. A review of physician turnover: rates, causes, and consequences. *American Journal of Medical Quality*, 19(2):56–66, 2004.

[74] Daniel R. Murphy, Ashley N.D. Meyer, Elise Russo, Dean F. Sittig, Li Wei, and Hardeep Singh. The Burden of Inbox Notifications in Commercial Electronic Health Records. *JAMA internal medicine*, 176(4):559–560, April 2016.

[75] Daniel R. Murphy, Brian Reis, Himabindu Kadiyala, Kamal Hirani, Dean F. Sittig, Myrna M. Khan, and Hardeep Singh. Electronic health record-based messages to primary care providers: valuable information or just noise? *Archives of Internal Medicine*, 172(3):283–285, February 2012.

[76] Daniel R. Murphy, Brian Reis, Dean F. Sittig, and Hardeep Singh. Notifications received by primary care practitioners in electronic health records: a taxonomy

and time analysis. *The American Journal of Medicine*, 125(2):209.e1–7, February 2012.

[77] Tore Opsahl, Filip Agneessens, and John Skvoretz. Node centrality in weighted networks: Generalizing degree and shortest paths. *Social Networks*, 32(3):245–251, July 2010.

[78] Michael R Oreskovich, Tait Shanafelt, Lotte N Dyrbye, Litjen Tan, Wayne Sotile, Daniel Satele, Colin P West, Jeff Sloan, and Sonja Boone. The prevalence of substance use disorders in american physicians. *The American journal on addictions*, 24(1):30–38, 2015.

[79] Sandhya Rao, Timothy G Ferris, Michael K Hidrue, Sara R Lehrhoff, Sara Lenz, James Heffernan, Kathleen E McKee, and Marcela G Del Carmen. Physician burnout, engagement and career satisfaction in a large academic medical practice. *Clinical medicine & research*, 18(1):3–10, 2020.

[80] Miriam Reisman. EHRs: the challenge of making electronic data usable and interoperable. *Pharmacy and Therapeutics*, 42(9):572–575, 2017.

[81] Pál Révész. *Random walk in random and non-random environments*. World Scientific, 2005.

[82] Wendi Rieb. *Increasing patient throughput in the MGH Cancer Center Infusion Unit*. PhD thesis, Massachusetts Institute of Technology, 01 2015.

[83] Michelle P Salyers, Kelsey A Bonfils, Lauren Luther, Ruth L Firmin, Dominique A White, Erin L Adams, and Angela L Rollins. The relationship between professional burnout and quality and safety in healthcare: a meta-analysis. *Journal of general internal medicine*, 32(4):475–482, 2017.

[84] Wilmar B Schaufeli and Arnold B Bakker. Utrecht work engagement scale: Preliminary manual. *Occupational Health Psychology Unit, Utrecht University, Utrecht*, 26(1):64–100, 2003.

[85] John Scott. Social network analysis. *Sociology*, 22(1):109–127, 1988.

[86] Piia Seppälä, Saija Mauno, Taru Feldt, Jari Hakanen, Ulla Kinnunen, Asko Tolvanen, and Wilmar Schaufeli. The Construct Validity of the Utrecht Work Engagement Scale: Multisample and Longitudinal Evidence. *Journal of Happiness Studies*, 10(4):459, May 2008.

[87] Suleyman Sevinc, Ulus Ali Sanli, and Erdem Goker. Algorithms for scheduling of chemotherapy plans. *Computers in biology and medicine*, 43(12):2103–2109, 2013.

[88] Tait Shanafelt, Joel Goh, and Christine Sinsky. The business case for investing in physician well-being. *JAMA internal medicine*, 177(12):1826–1832, 2017.

[89] Tait D. Shanafelt. Enhancing meaning in work: a prescription for preventing physician burnout and promoting patient-centered care. *JAMA*, 302(12):1338–1340, September 2009.

[90] Tait D. Shanafelt, Sonja Boone, Litjen Tan, Lotte N. Dyrbye, Wayne Sotile, Daniel Satele, Colin P. West, Jeff Sloan, and Michael R. Oreskovich. Burnout and satisfaction with work-life balance among US physicians relative to the general US population. *Archives of Internal Medicine*, 172(18):1377–1385, October 2012.

[91] Tait D. Shanafelt, Lotte N. Dyrbye, Christine Sinsky, Omar Hasan, Daniel Satele, Jeff Sloan, and Colin P. West. Relationship Between Clerical Burden and Characteristics of the Electronic Environment With Physician Burnout and Professional Satisfaction. *Mayo Clinic Proceedings*, 91(7):836–848, July 2016.

[92] Tait D. Shanafelt, Omar Hasan, Lotte N. Dyrbye, Christine Sinsky, Daniel Satele, Jeff Sloan, and Colin P. West. Changes in Burnout and Satisfaction With Work-Life Balance in Physicians and the General US Working Population Between 2011 and 2014. *Mayo Clinic Proceedings*, 90(12):1600–1613, December 2015.

[93] Tait D Shanafelt, Michelle Mungo, Jaime Schmitgen, Kristin A Storz, David Reeves, Sharonne N Hayes, Jeff A Sloan, Stephen J Swensen, and Steven J Buskirk. Longitudinal study evaluating the association between physician burnout and changes in professional work effort. In *Mayo Clinic Proceedings*, volume 91, pages 422–431. Elsevier, 2016.

[94] Tait D. Shanafelt and John H. Noseworthy. Executive Leadership and Physician Well-being: Nine Organizational Strategies to Promote Engagement and Reduce Burnout. *Mayo Clinic Proceedings*, 92(1):129–146, January 2017.

[95] Tait D. Shanafelt, Jeff A. Sloan, and Thomas M. Habermann. The well-being of physicians. *The American Journal of Medicine*, 114(6):513–519, April 2003.

[96] Tait D Shanafelt, Colin P West, Christine Sinsky, Mickey Trockel, Michael Tutty, Daniel V Satele, Lindsey E Carlasare, and Lotte N Dyrbye. Changes in burnout and satisfaction with work-life integration in physicians and the general us working population between 2011 and 2017. In *Mayo Clinic Proceedings*, volume 94, pages 1681–1694. Elsevier, 2019.

[97] Scott A Shipman and Christine A Sinsky. Expanding primary care capacity by reducing waste and improving the efficiency of care. *Health Affairs*, 32(11):1990–1997, 2013.

[98] Christine Sinsky, Lacey Colligan, Ling Li, Mirela Prgomet, Sam Reynolds, Lindsey Goeders, Johanna Westbrook, Michael Tutty, and George Blike. Allocation of Physician Time in Ambulatory Practice: A Time and Motion Study in 4 Specialties. *Annals of Internal Medicine*, 165(11):753–760, December 2016.

[99] Christine A. Sinsky, Adam Rule, Genna Cohen, Brian G. Arndt, Tait D. Shanafelt, Christopher D. Sharp, Sally L. Baxter, Ming Tai-Seale, Sherry Yan, You Chen, Julia Adler-Milstein, and Michelle Hribar. Metrics for assessing physician activity using electronic health record log data. *Journal of the American Medical Informatics Association: JAMIA*, 27(4):639–643, April 2020.

[100] Christine A. Sinsky, Rachel Willard-Grace, Andrew M. Schutzbank, Thomas A. Sinsky, David Margolius, and Thomas Bodenheimer. In search of joy in practice: a report of 23 high-functioning primary care practices. *Annals of Family Medicine*, 11(3):272–278, June 2013.

[101] Harold C Sox. Leaving (internal) medicine. *Annals of Internal Medicine*, 144(1):57, 2006.

[102] Ming Tai-Seale, Cliff W. Olson, Jinnan Li, Albert S. Chan, Criss Morikawa, Meg Durbin, Wei Wang, and Harold S. Luft. Electronic Health Record Logs Indicate That Physicians Split Time Evenly Between Seeing Patients And Desktop Medicine. *Health Affairs (Project Hope)*, 36(4):655–662, April 2017.

[103] John Thune, Lamar Alexander, Pat Roberts, Richard Burr, and Mike Enzi. Where is HITECH's $35 billion dollar investment going? *Health Affairs Blog*, 2015.

[104] Ayten Turkcan, Bo Zeng, and Mark Lawley. Chemotherapy operations planning and scheduling. *IIE Transactions on Healthcare Systems Engineering*, 2(1):31–49, 2012.

[105] André van Vliet. An improved lower bound for on-line bin packing algorithms. *Information processing letters*, 43(5):277–284, 1992.

[106] Jean E. Wallace, Jane B. Lemaire, and William A. Ghali. Physician wellness: a missing quality indicator. *Lancet (London, England)*, 374(9702):1714–1721, November 2009.

[107] Colin P. West, Liselotte N. Dyrbye, Patricia J. Erwin, and Tait D. Shanafelt. Interventions to prevent and reduce physician burnout: a systematic review and meta-analysis. *Lancet (London, England)*, 388(10057):2272–2281, November 2016.

[108] E. S. Williams, T. R. Konrad, W. E. Scheckler, D. E. Pathman, M. Linzer, J. E. McMurray, M. Gerrity, and M. Schwartz. Understanding physicians' intentions to withdraw from practice: the role of job satisfaction, job stress, mental and physical health. *Health Care Management Review*, 26(1):7–19, 2001.

[109] Jonathan C Woodall, Tracy Gosselin, Amy Boswell, Michael Murr, and Brian T Denton. Improving patient access to chemotherapy treatment at duke cancer institute. *Interfaces*, 43(5):449–461, 2013.

[110] Richard A. Young, Sandra K. Burge, Kaparaboyna A. Kumar, Jocelyn M. Wilson, and Daniela F. Ortiz. A Time-Motion Study of Primary Care Physicians' Work in the Electronic Health Record Era. *Family Medicine*, 50(2):91–99, February 2018.

[111] Ji Zhu and Trevor Hastie. Classification of gene microarrays by penalized logistic regression. *Biostatistics*, 5(3):427–443, 2004.