Transactions of the SDPS:
Journal of Integrated Design and Process Science
XX (XXXX) XX-XX
DOI 10.3233/jid-201x-xxxx
http://www.sdpsnet.org

Agent-Based System Design for Service Process Scheduling: Challenges, Approaches and Opportunities

Farnaz Dargahi, Chun Wang*, Mohammad F. H. Bhuiyan and Hamidreza Mehrizi

Concordia Institute for Information Systems Engineering, Concordia University, Montreal, Canada

Abstract Compared with traditional manufacturing scheduling, service process scheduling poses additional challenges attributable to the significant customer involvement in service processes. In services, there are typically no inventoried products, which make the service provider's capacity more sensitive to dynamic changes. Service process scheduling objectives are also more complicated due to the consideration of customer preferences, customer waiting costs and human resource costs. After describing the Unified Services Theory and analysing its scheduling implications, this paper reviews the research literature on service process scheduling system design with a particular emphasis on agent-based approaches. Major issues in agent-based service process scheduling systems design are discussed and research opportunities are identified. The survey of the literature reveals that despite of many domain-specific designs in agent-based service process scheduling, there is a lack of general problem formulations, classifications, solution frameworks, and test beds. Constructing these general models for service process scheduling system design will facilitate the collaboration of researchers in this area and guide the effective development of integrated service process scheduling systems.

Keywords: Services, agent-based systems, decentralized scheduling, dynamic scheduling, auctions

1 Introduction

Scheduling is a decision-making process which allocates limited resources to tasks over time while satisfying certain constraints and optimizing one or more objectives. Scheduling problems are common to many domains such as manufacturing and services. The number and variety of scheduling problem models is astounding. In spite of the various presentations, most of the models can fit into a four-element structure which consists of activities, resources, constraints, and objectives (Wang, 2007). Using the four elements, Wall (1996) defines general resource constrained scheduling problems as given a set of activities that must be executed, a set of resources with which to perform the activities, a set of constraints which must be satisfied, and a set of objectives with which to judge a schedule's performance, finding the best way to assign the resources to the activities at specific times such that all of the constraints are satisfied and the best objective measures are produced.

The scheduling problems in service settings can be somewhat different from those in manufacturing. As summarized in Pinedo (2009), in manufacturing an activity usually transforms a physical component and adds value to it; resources are typically referred to as machines and the configuration of machines;

* Corresponding author. Email: chun.wang@concordia.ca Tel: (+1)514-8482424 ext. 5628.

 Published by IOS Press

objectives are typically a function of the completion times, the due dates, and the deadlines of the jobs. In service settings an activity usually involves people. It can be, for example, a meeting that has to be attended by certain people, a flight that transports passengers, an operation that has to be done by a surgeon on a given day. Services usually require both physical and human resources. In addition, the operational constraints in services can take diverse forms. A typical type is capacity requirements. They are important in reservation systems, in timetabling of meetings as well as in transportation planning and scheduling. In service settings, additional factors such as personnel costs, customer waiting costs and customer preferences are often considered in the objective function.

The differences between manufacturing and service process scheduling are mainly derived from the fundamental characteristic which defines service processes. A service significantly involves customer inputs (Sampson & Froehle, 2006). In other words, in order for a service to be produced, a customer has to present personally or he/she has to present his/her belongings or information. Compared with classical manufacturing scheduling models, this significant involvement of customer inputs presents additional challenges including distributed and dynamic environments, the presence of private customer information and often considerably more complicated scheduling objectives (we will explain these challenges in details in the next section).

The objective of this paper is not to provide an extensive survey of general service process scheduling models, but to focus on the models that take an agent-oriented paradigm which, we believe, is suitable for tackling service process scheduling challenges given its strength on dealing with distributed, dynamic and complex environments. An earlier survey of multi-agent systems for manufacturing process planning and scheduling can be found in Shen et al. (2006). Detailed descriptions of classical service process scheduling models can be found in Pinedo (2009).

The rest of the paper is organized as follows. In Section 2, we first describe the Unified Services Theory (Sampson, 2001), which categorically defines services. We then analyze the challenges in service process scheduling system design in light of the theory. In Section 3, we provide a brief overview of traditional approaches to service process scheduling system design. In Section 4, we review literature on agent-based service process scheduling system design. Major design issues and research opportunities are discussed in Section 5. Section 6 concludes the paper.

2 Unified Services Theory and Its Scheduling Implications

Services have been commonly defined as intangible products (Pearce, 1981, p. 390; Bannock et al., 1982, p. 372; Harvey, 1998, p. 596). In other words, a service typically does not result in the ownership of anything (Kotler & Keller, 2006, p. 402). Intangibility is an important characteristic of services. However, as stated in Sampson and Froehle (2006), it does not serve as a sufficient condition which defines a production process as a service. For example, software development results in a product that is intangible (computer code), but the output can indeed be inventoried and used or sold later. Unified Services Theory, on the other hand, identifies a single commonality that comprises all services. It defines what services are and what they are not. To facilitate the analysis of service implications to scheduling, it is useful to first introduce the Unified Service Theory.

2.1 Unified services theory

The Unified Services Theory (UST) is formally stated as follows (Sampson, 2001, p. 16):

"With service processes, the customer provides significant inputs into the production process. With manufacturing processes, groups of customers may contribute ideas to the design of the product, but individual customers' only participation is to select and consume the output. All managerial themes unique to services are founded in this distinction."

The most important component in UST is customer inputs which distinguish services from manufacturing processes and are the root cause of the unique issues and challenges of services management. The literature has typically identified three general types of customer inputs (Wemmerlov, 1990): the customer's self, his belongings or other tangible objects and information. Customer-self inputs are common in services involving co-production (i.e., the employment of customer labor in the process) and in services involving the physical presence of the customer. Typical examples are health care offices, buffet restaurants and taxi services. These service providers can prepare for production, but they cannot execute the actual service process until necessary customer-self inputs are present. Tangible belongings (or property) and physical objects make up another type of input a customer can provide to the service process. One's car is an essential input into the automobile repair service process and one's clothing is a necessary input to the dry cleaning service process. Providing tangible inputs often allows the service process to proceed even without the customer being physically present. Customer-provided information is a third type of input to the service process. For example, the tax return preparation process requires that customers provide financial information as process inputs. Without that information input the service production process cannot begin.

The UST reveals principles that are common to the wide range of services and provides a unifying foundation for various theories and models of service operations. As demonstrated in Sampson and Froehle (2006), the UST has significant operational corollaries pertaining services management processes. Among them, capacity management and demand management significantly rely on the scheduling of service resources. In the rest of this section, we analyze the implications of UST to service process scheduling. We also present challenges in designing service process scheduling systems.

2.2 Service process scheduling implications

Scheduling plays an important role in service management due to the perishable nature of service provider's capacity. A service provider has to pay scheduled workers even though there are no customers currently needing services. In other words, the service provider's capacity to produce the service is time-sensitive and cannot be inventorized by producing to stock. This high "operating leverage" implies that many service operations will be much more cost-competitive if the service providers effectively manage variable demand (Hur et al., 2004; Jack & Powers, 2004), which gives them higher utilization levels (Sampson, 2001, p. 240) or, alternately, manage capacity, which increase their volumes.

The management of demand and capacity involves the allocation of service orders and resources over time, which is essentially a scheduling activity. On the demand management side, reservation systems schedule customer inputs into the production process such that waiting times are minimized. On the capacity management side, service managers schedule full- and part-time personnel to meet the expected workload for a future day. When the day of service arrives, if a significant gap is present between the experienced workload so far and the scheduled staff capacity, service managers will attempt to make an immediate adjustment to the staff schedule by changing station assignment, shifting breaks, or calling in additional workers (Hur et al., 2004). Compared with classical manufacturing scheduling, service process scheduling presents different challenges attributable to significant customer inputs in service production processes. In the following, we describe three important service process scheduling challenges, namely distributed and dynamic environments, complicated objectives and customers' private information.

2.2.1 Distributed and dynamic environment

The requirement of customer inputs in services leads to a distributed and dynamic scheduling environment. First, the information needed for computing schedules, e.g. customers' availability and preference information, is scattered among possibly a large number of customers. Collecting the information and keep it up to date can be challenging tasks. Secondly, service process scheduling has to be robust in accommodating contingencies caused by the customer involvement in service production. Uncertainty in customer demand, resource availability, service times, customer cancelations and no-shows make the scheduling of services a complex dynamic process. Customers may ask to include additional tasks that are not anticipated, or to adapt to changes to several tasks, or to neglect certain tasks. The resources available for performing tasks are subject to changes as well. Certain resources can become unavailable, and additional resources may need to be introduced. The beginning time and the processing time of a task are also subject to variations. A task can take more or less time than anticipated, and the customer inputs can arrive early or late. An optimal schedule, generated after considerable effort, may rapidly become unacceptable because of unforeseen dynamic situations. Since service capacity cannot be inventorized by producing to stock, customers who fail to present their inputs according to the schedule can lead to poor resource utilization, lower revenues and longer waiting times. The time-sensitive nature of service capacities signifies the need for more robust dynamic scheduling approaches. In addition, unlike the manufacturing environments where the number of resources (which are typically machines) is usually fixed (at least for the short term), in services, the number of resources (e.g. people, rooms, and trucks) may vary over time.

The service process scheduling is further complicated by the fact that customers' needs for services have varying degrees of urgency, and some decisions about non-urgent requests must be made in advance of having complete information about urgent and emergency demands. Take patient scheduling in diagnostic services, such as magnetic resonance imaging (MRI) scanning or computed tomography (CT) scanning, as an example. The low-priority demand (outpatients) must be booked (often weeks in advance) before knowing the highly unpredictable high-priority demand (inpatients). To accommodate the demand imposed by the highly dynamic high priority inpatients, the hospital is forced to reserve a significant portion of the total capacity for this unknown high-priority demand leaving little room for outpatients. This results in unused capacity on days when inpatient demand is lower than expected and thus longer waiting times for outpatients than might be the case if this unused capacity could be utilized.

2.2.2 ***Complicated objectives***

Planning and scheduling objectives in service industries are often considerably more complicated than those in manufacturing. Scheduling objectives in manufacturing are typically a function of the completion times, the due dates, and the deadlines of the jobs. Objectives in services may have additional dimensions. In contrast to manufacturing, the number of resources in a service environment may be variable (e.g. the number of full-time and part-time people employed). Because of this, there may be a different type of objective that tries to minimize the number of resources used and/or minimize the cost associated with the use of these resources. This is a typical objective of capacity management. In addition, customer preferences regarding the timing of delivering their inputs should also be considered in service process scheduling as they represent customer values over a schedule. For example, in healthcare services, patients want more personalized care, which includes involvement in selecting appointment-times. Some patients prefer an appointment on the day they call, or soon thereafter, and the day of the week or the time of the appointment is not particularly important to them. Others prefer a particular day of week and a convenient time. They do not mind waiting for convenience. In both private and public healthcare systems, healthcare managers care about having high scores on patient satisfaction surveys. In addition, offering patients a convenient appointment time can decrease the number of no-shows and thereby increase operational efficiency (Wang and Gupta, 2011).

2.2.3 ***Customers' private information***

Service processes involve significant customer inputs, which, in many cases, require that services are produced and consumed at the same time. Scheduling systems are used to synchronize the timing of the use of the different types of resources and the presence of customer inputs. To compute optimal schedules, ideally, the scheduler should know the complete customer availability information within the scheduling horizon. However, collecting the availability information across a large number of customers requires a significant amount of communication between the scheduler and the customers. This amount of communication can incur high administrative costs if the collecting procedure is not automated, which is the case of most existing service process scheduling systems. The issue is further complicated by the fact that customers are reluctant to reveal their complete availability because they treat their personal schedule as their private information. They are actually motivated to protect their privacy. Therefore, service process scheduling systems should also be designed in a way that they are able to elicit necessary customer availability information to compute high quality schedules. The computation spent on eliciting customer's availability information is referred to as elicitation complexity of the system. Elicitation complexity is imposed by the privacy constraint of the customers and calls for game theoretic approaches.

3 Centralized Service process scheduling Approaches

Traditional service process scheduling approaches usually assume a centralized environment in which a scheduler has all needed information to compute the schedule. Various service process scheduling models have been proposed, implemented, and evaluated for several decades. Generally speaking, the solution methods form two distinct classes: exact methods and heuristic methods. Exact methods are guaranteed to find a solution if it exists, and typically provide some indication if no solution can be found. However, given the NP-hard nature of service process scheduling models, exact methods are not practical for non-trivial problem instances. Heuristic methods do not guarantee optimization, but typically assure experimentally or analytically some degree of optimality in their solutions. They are usually quick and are practical ways of solving larger size scheduling problems. In this section, we briefly review some general heuristic methods and their application to service scheduling.

3.1 Genetic algorithms

Genetic Algorithms (GAs) are a set of global search and optimization methods for solving complex optimization problems with a large search space. With the objective of reaching the "best" solution, GAs systematically evolve a population of candidate solutions by using evolutionary computational processes inspired by genetic variation and natural selection. One of the earliest GAs for scheduling was proposed by Davis (1985). In his paper, Davis suggested an indirect representation which can be decoded to form the actual schedule of the scheduling problem. GAs have been applied to many service scheduling problems. For example, Ghaemi et al. (2007) proposed co-evaluation algorithm for university timetabling problem. Paechter et al. (1995, 1996) applied memetic algorithm for course timetabling. The memetic algorithm explorer the neighbourhood of the solution obtained by GA and navigates the search towards the local optima. Graph colouring heuristics are used by Burke et al. (1995, 1996, & 1998) to improve and accelerate the search process in timetabling. Burke et al. (1995) also developed a hybrid GA to ensure the most fundamental constraints are never violated in timetabling problem. They showed that the algorithm is guaranteed to produce a feasible solution by hard coding constraints and using hybrid crossover operator. In addition to timetabling, GAs have also been used to solve the scheduling problems in healthcare, such as patient scheduling and nurse scheduling (Petrovic et al., 2011; Aickelin & Dowsland, 2001).

3.2 Simulated annealing

Simulated Annealing (SA), is a neighbourhood search method. Rather than always choosing the direction of the best improvement, which gives steepest-ascent hill-climbing, SA initially chooses random or semi-random direction but over time comes to prefer the direction of the best improvement. The direction selection process is controlled by some sort of temporal parameter, which is usually called 'temperature' by analogy with real annealing. SA approaches require a schedule representation as well as a neighbourhood operator for moving from the current solution to a candidate solution. Annealing methods allow jumps to worse solutions and thus often avoid local sub-optimal solutions (Kirkpatrick et al., 1983). Quality of solutions produced by a SA implementation depends on the correct choice of solution space and neighbourhood, as well as the parameters that govern the cooling schedule. SA has been applied to service scheduling. For example, Gunawan et al. (2007) used a hybrid algorithm which consists of an integer programming, a greedy heuristic and a modified SA algorithm for solving large scale timetabling problems. Bailey et al. (1997) solved a nurse scheduling problem using SA and compared its performance with integer programming and a GA. They found that, for a given quality, their algorithm was faster than the GA and integer programming for the set of nurse scheduling testing problems.

3.3 Tabu search

Tabu search (TS) is similar to SA in that it also moves from one schedule to another with the next schedule being possibly worse than the one before. The difference is in the mechanism by which moves to new schedules are accepted. A TS maintains a list of tabu moves, representing schedules which, having been visited recently, are forbidden in order to diversify the directions in which search proceeds. TS has been proposed to compute high complexity large size health care service scheduling. Dowsland (1998) used tabu search with strategic oscillation for nurse scheduling. The objective is to ensure adequate nurses are on duty at all times while incorporating individual preferences and requests for days off in a way that is seen to be fair to all nurses. The method uses a variant of TS which oscillates between solutions with feasible nurse coverage and then applies nurse preferences to improve upon the solution. Demeester et al. (2010) proposed a hybrid TS algorithm for patient admission scheduling. It automatically assigns patients to beds in the appropriate departments by considering medical needs of the patients as well as their preferences while keeping the number of patients in the different departments balanced. The method uses a TS algorithm hybridized with a token-ring and a variable neighbourhood descent algorithm. To university course timetabling problems, TS has also been applied (Hertz, 1991; Hertz, 1992).

3.4 Constraint logic programming

Many service scheduling problems can be modelled as constraint satisfaction problems (CSP). In a CSP, values which satisfy a set of constraints must be found for a set of discrete variables with finite domains. Constraint satisfaction is a search procedure that operates in the space of constraint sets rather than in that of the solution sets. A Constraint Logic Programming (CLP) provides the ability to declare variables and their domains for CSP problems. Examples of applying CLP to service scheduling problems can be found in Gueret et al. (1995), Henz and Wurtz (1995), and Abdennadher and Schlenker (1999).

3.5 Approaches considering customer preferences and dynamic environment

Because of the computational complexity involved in creating schedules that simultaneously consider customer preferences and scheduling objectives, a limited research in centralized service scheduling considered customer preferences. Wang and Gupta (2011) proposed a heuristic approach for patient scheduling which captures customer preferences. The method has two components. The first one dynamically learns patient’s preferences, updates estimate of acceptance probabilities. The second one uses the acceptance probability information for booking decisions. Jaumard et al. (1998) proposed an integer programming model accommodating workers’ preferences. The problem was solved using Dantzig-Wolfe decomposition. The objective was to minimize salary costs and maximize nurse preferences. Azaiez and Sharif (2005) developed a 0-1 linear goal programming model for the nurse scheduling in a hospital in Saudi Arabia. Nurse’s preferences for shift time are obtained from a survey consisting of 15 multiple choices. Nurses’ preferences were combined with hospital constraints to develop the linear goal programming model.

Centralized service scheduling usually deal with dynamic environment using simulation based approaches. A simulation is the imitation of the operation of a real-world process or system over time (Groothuis & Merode, 2001). An advantage of simulation study over heuristic approaches is the ability of modelling complex systems and representing environmental variables. Hancock and Walter (1984) conducted a simulation study based on historical data of patient arrival. The simulation is used to determine the number of procedures that would be performed in each day of the week. Groothuis and Merode (2001) applied discrete event simulation technique to optimize the use of catheterization capacity in a hospital. Ho and Lau (1999) proposed a simulation based method for evaluating the impact of different combinations of the dynamic environmental factors such as no-shows, service times, and the number of customers per service session to the quality of service schedules.

The above mentioned traditional scheduling methods encounter great difficulties when they are applied to real-world situations. This is because they use simplified theoretical models and are essentially centralized in the sense that all computations are carried out in a central computing unit. The intelligent agent technologies, on the other hand, suggest an innovative, lightweight approach to scheduling problems. The main characteristic of intelligent agents is their autonomy. Each agent makes its own decisions, based on its internal state and on the information it receives from its environment; so each agent can keep its independency from the rest of system. In other words each agent according to its private information may use different policy independently from the rest of the system. Agent-based systems are inherently distributed and robust in dynamic environments. Agents can retrieve information from different resources, analyze them, filtering redundant information, select and present the data by an interface which is interested by users. Another feature of agents is their sociability. Agents can communicate with each other and exchange any kind of information. By this way they can overcomes inconsistency among their local schedules and resolve errors and collaborate in the process of scheduling. Thus according to the properties of agent-based systems, agent-based approach can be a good candidate for service scheduling.

4 Literature on Agent-Based Service Scheduling System Design

Agent-based service scheduling system design is essentially a distributed approach which is more flexible, efficient, and adaptable to real-world dynamic environments (Shen et al., 2006). By applying agent-based service scheduling system architecture, the distributed nature of service scheduling is naturally modelled. In addition, each agent can be assigned different objectives. In this way, the complicated multiple objectives in service scheduling can be decomposed to individual agents. This decomposition significantly simplifies the modelling of the objectives (Jennings, 2001). Agent-based scheduling systems have been proposed for several important service sectors. However, there is a lack of general problem formulations, classifications, solution frameworks, and test beds in service scheduling. We therefore take a domain specific approach. The service process scheduling literature has concentrated on several representative domains such as meeting, healthcare, transportation, and computing services. We review these application domains through the lens of how agent-based system design approach addresses service process scheduling challenges. Since the challenges of distributed scheduling information and complicated multiple objectives have been naturally modelled in agent-oriented design paradigm, in this section, we focus on how agent-based scheduling system design tackles the challenges of dynamic environment and users' private information.

4.1 Meeting scheduling

Meeting scheduling problem signifies a decision-making process affecting several users, in which it is necessary to decide ”when” and ”where”, one or more meetings should be scheduled (Hassine et al., 2004). Since it usually involves inputs of multiple users, meeting scheduling can be classified as a service scheduling problem. Agent-based meeting scheduling approaches have been proposed in the literature. Some of them are distributed implementation of constraint satisfaction algorithms in the multiagent systems environment. In the multiagent meeting scheduling system developed by Franzin et al. (2002), agents communicate in several proposal phases. Whenever agents communicate during the proposal phases, the information they exchange can be used to build an approximation of the constraint set of the other agents. In other words each agent in the proposal phase is able to elicit other agent’s availability. To deal with the challenge of dynamic environment, Hassine et al. (2004) formalize meeting scheduling as a dynamic valued constraint satisfaction problem. Agents negotiate with each other to achieve a schedule in a way that maximizes global utility. In the negotiation process host agent proposes a set of timeslots as a solution to the other agents who participate in the meeting. Each participant agent that has received this message ranks the obtained time slots according to its preferences and constraints and returns them to the proposer agent. Proposer agent tries to find the best solution, which maximizes its utility, from the received time slots. The same process resumes until an agreement is reached among all of the agents. Course timetabling at universities, which can be seen as a type of meeting scheduling problem, is also modeled as a constraint satisfaction problem by Meisels and Kaplansky (2003). Inter agent negotiation protocol is used to overcome inconsistency among local schedules.

The presence of users' private information is also addressed in agent-based meeting scheduling. Wainer et al. (2007) defined four levels of privacy protocol (or modes of agents' interaction) to model users' private information, namely, full information protocol, approval protocol, voting protocol and suggestion protocol. These modes of interaction are defined based on whether the participants are comfortable in sharing their private information with the host or not during the negotiation process. In Modi et al. (2004), agents' private information is modelled as their utilities. Each agent makes a decision about accepting a meeting time based on how the decision will impact its utility. The utility of a timeslot is calculated based on the difference between the value of meeting scheduled in the timeslot and the predicted cost of negotiating with other agents. Crawford and Veloso (2004) designed a mechanism for meeting scheduling which is incentive compatible. A mechanism is incentive compatible if it is every agent's dominant strategy to reveal their private utility values truthfully. The mechanism motivates agents to reveal their valuation for each of the feasible schedules. The schedule that maximizes the social welfare is selected. Agent's payments are VCG auction payments which justifies the incentive compatibility of the mechanism. Iterative auction are also used in agent-based meeting scheduling. In a course timetabling system proposed by Sönmez and Ünver (2007), students are assigned certain amount of bid endowments and they bid for different schedules of courses using the endowments assigned. Students are modelled as price-takers under a belief system. In other words students' bids are based on their guess about the market-clearing price they will face. Krishna and Ünver (2007) also proposed a course bidding system and conducted a field test at the Ross School of Business, University of Michigan, in spring 2004 semester. In their biding system student bids are used to infer students' preferences over courses and to determine their priorities for courses. In addition to users' private information, the challenge of dynamic environment is also addressed in agent-based meeting scheduling. Typical examples include Wainer et al. (2007), Modi et al. (2004) and Sönmez and Ünver (2007).

4.2 Healthcare

Agent-based approach in which patients and hospital resources are modelled as autonomous agents with their own goals, reflects the decentralized structures of health care environment. Most of the agent-based healthcare scheduling literature focuses on the challenge of distributed and dynamic environment of healthcare management. In a recent research on operation rooms scheduling, Zhiming (2011) developed a two stage approach which addresses the challenges of dynamic scheduling. Mixed integer programming is used in the first stage for assigning surgical operation to each operation room. The second stage utilizes a dynamic rescheduling approach, in which agents reallocate tasks among them using the contract net protocol in a way that minimize the cost of the operation rooms.

Agent-based approaches are also proposed for patient scheduling. Hannebauer and Muller (2001) formulated patient scheduling as a distributed constraint optimization problem. They proposed the Multi-phase Agreement Finding (MPAF) algorithm for coordinating the agents and covering the constraints. MPAF consists of two phases, the proposal phase and the assignment phase. In the proposal phase diagnostic unit agent selects a set of feasible appointment timeslots based on its optimization criteria and proposes to the patient agent. In the assignment phase, the patient agent decides whether to accept the proposed timeslots. This decision is made based on the agent's scheduling constraints and its scheduling objective which is to minimize the waiting time between appointments. Other agent-based patient scheduling approaches model the scheduling environment as a market. Given the distributed and dynamic nature of patient scheduling, markets can efficiently distribute scare resources between patients. Paulussen et al. (2003) developed a bidding mechanism for patient scheduling, in which patient agents communicate their (private) utility for certain time-slots on a resource via a price mechanism. The price that patient agents are willing to pay is the difference between the cost-value of the current allocation and the cost-value for the wanted appointment. Resources are assigned to the patients that are willing to pay the highest price (to the patients who gain the highest health sate improvement). The scheduling objective is to maximize resource utilization and minimize patient stay time in hospital. For patients who need to schedule several related appointments, a multi-round auction mechanism is proposed by Hosseini et al. (2011). In this approach, patients calculate the value of obtaining each resource by solving their Markov decision problem. In each round of auction, agents submit their bids; auctioneer determines the winner and moves to the next step. The objective of winner determination is to minimize the global regret values of patients. Regret value of a patient on a resource is defined as the difference in value between getting the resource and not getting the resource given patient's current health state.

Agent-based approaches are also proposed for nurse timetabling. Grano et al. (2009) proposed an auction based nurse scheduling approach that considers both nurse preferences and hospital requirements. In the auction nurses bid for work shifts and rest day using the points instead of money value. So in the bidding stage nurse's private information which consists of availability and preferences for specific days and shifts are obtained. Winners are selected using an optimization model which seeks to award shifts to the highest bidders while simultaneously meeting hospital requirements.

4.3 Transportation services

Agent-based approach has been adopted in transportation planning and scheduling research for more than two decades. Fischer et al. (1995) pointed out that transportation planning and scheduling are inherently distributed, complex tasks. Geographically, trucks and jobs are distributed and also maintain some level of autonomy. To implement traditional methods, a scheduler must gather a large amount of information to a central place where the solution can be computed. However, using agent-based approach, an agent only requires local information. In their review on multiagent systems in logistics, Lang et al. (2008) concluded that planning and scheduling problems in transportation have specifications that comply with particular capabilities of agent systems. Specifically, these systems are able to deal with inter-organizational and event driven scheduling settings that meet supply chain's planning and execution requirements. Davidsson et al. (2005) also identified a number of positive aspects of the agent-based approaches to logistics. Existing surveys (Lang et al., 2008; Davidsson et al., 2005) mainly focus the research addressing the distributed and dynamic aspects of transportation services. In the rest of this section, we review papers focusing on the challenge of the presence of customers' private information, which is mainly tackled by the design of various auction systems in the context of multiagent systems.

Auction mechanisms, especially combinatorial auctions, have been adopted by a large number of shippers and 3PL (third party logistic) providers. Leading companies such as Wal-Mart, Procter & Gamble and Sears have used combinatorial auctions to reduce their logistic costs (Sheffi, 2004). Song and Regan (2003) proposed an auction based mechanism, the Collaborative Carrier Network, for carriers to exchange their excess capacities in a TL (truckload) spot-market. Through this network, carriers can buy and sell transportation capacities. The network is structured as a group of auctions launched by carriers. Each carrier can be both a contractor and a sub-contractor in different auctions. A carrier will launch at most one auction at a time and that if new loads come in during the previous auction round, they will be simply held and wait for the next round. The network attempts to ease the exchange of information, drop transaction cost and make it possible for both carriers and shippers to access larger markets. Kwon et al. (2005) also proposed an iterative auction mechanism for TL transportation procurement. Each agent (carrier) bids for a package of lanes. A descending multi-round format is used to allocate lane packages to the agents. First, agents compute their preferred packages based on their cost structures and submit them to the auctioneer. Then the auctioneer performs a provisional allocation of lanes to the agents by solving a winner determination problem (WD) with objective of minimizing the payments. Simulation results showed that both carriers and shippers reduced their cost through a better collaboration. For the LTL (less than truckload) setting, Krajewska and Kopfer (2006b) proposed an auction model for the collaboration among individual freight forwarding entities. Cooperating forwarders exchange their orders through a combinatorial auction. The auction is individually rational, which means each individual partner increase its profit by participating in the coalition.

Effective collaboration among agents in a distributed system leads to better utilization of resources and, thus, greater efficiency and profit for the whole system. However, before entering into the partnership, agents have to agree upon how to share the profit resulted from the collaboration. In a collaborative environment where, for example, carrier companies belong to a common holding organization, profit sharing may not require incentive compatible mechanisms. Gujo et al. (2009) proposed an exchange mechanism, called ComEx, for inter-enterprise logistic services. In ComEx, transportation capacity in each division is managed by a profit centre which can possibly exchange delivery orders with other profit centres based on the geographical zones and time windows of the orders. The gained profit is shared proportionally among profit centres based on the cost saving of each profit centers participating the exchange. A precondition of this type of profit sharing is that ComEx has access to the cost saving data of profit centers. ComEx works well in the collaborative setting. However it is not suitable for game theoretic settings where profit centres do not belong to a common holding organization and they may be reluctant to share their cost saving data. In this case, profit distribution mechanism based on game theory and combinatorial auction should be applied (Krajewska and Kopfer, 2006b; Gomber et al., 1997). Other agent-based models in transportation services distribute gained benefit of collaboration from a loss sharing rather than profit sharing perspective (Schönberger, 2005; Schönsleben & Hieber, 2004). Krajewska and Kopfer (2006a) present an overview of these benefit sharing models.

4.4 Computing services

Modern computing services aggregate a large number of independent computing and communication resources and data stores. They are built on the bases of distributed computing, grid computing and virtualization. Computing service environment is inherently complex, heterogeneous and dynamic. Service resource management systems need to provide mechanisms and tools that allow resource consumers (end users) and providers (resource owners) to express their requirements and facilitate the realization of their goals. This objective necessitates seamless scheduling of providers' resources to support dynamic scaling of users activities across multiple domains. Scheduling computing services under varying load, diverse application requirements and heterogeneous systems is a challenging problem. Agent-based approach can be an effective way to realize information sharing, unpredictable dynamism and increasing heterogeneity in computing service scheduling.

With the aim of tackling the challenge of dynamic environment in computing services, An et al. (2010) proposed a distributed negotiation mechanism for dynamic and uncertain resource demand and supply in computing as service (cloud computing) platform. The mechanism is an extension to alternating offers protocol with the feature of allowing agents to decommit from contracts at a cost. The mechanism facilitates the agents' negotiation over both a contract price and a decommitment penalty. They evaluated and compared their approach experimentally using representative scenarios and workloads, to both combinatorial auctions and the fixed-price model used by Amazon's EC2, and showed that their model achieves a higher social welfare. Scheduling mechanisms for computing services typically deal with the dynamics of both resource and service markets. Sim (2012) proposed a concurrent negotiation mechanism for agents to negotiate in multiple interrelated e-Markets. He developed an agent-based test bed consisting of provider agents and consumer agents acting on behalf of resource providers and consumers, respectively, and a set of broker agents. The mechanism consists of: (1) a bargaining-position-estimation strategy for the multilateral negotiations between consumer and broker agents in a service market and (2) a regression-based coordination strategy for concurrent negotiations between broker and provider agents in resource markets. The negotiation outcomes between broker and provider agents in a resource market can potentially influence the negotiation outcomes between broker and consumer agents in a service market. Using this mechanism, the broker agent accepts service requests from consumer agents, purchase resources from provider agents. The collection of resources which satisfy consumer agents' requirements is dynamically composed. Mobile agents are also designed for providing scalability in cloud computing. In Singh and Malhotra (2012), a mobile agent is capable of transporting its state from one environment to another with its data intact and performing appropriately in the new environment. The agents are supported with algorithms for searching another cloud with better response time when the approachable cloud becomes overloaded.

To deal with the challenge of customer's private information, game-theoretic based methods have been proposed to solve the resource allocation problem in network systems. Gagliano et al. (1995) presented an auction allocation of computing resources. In the proposed auction, computing tasks are provided sufficient intelligence to acquire resources by offering, bidding and exchanging them for funds. Wolski et al. (2001) compared commodities markets and auctions in grids in terms of price stability and market equilibrium. Zaman and Grosu (2011) studied and implemented combinatorial auction-based mechanisms for efficient provisioning and allocation of computing service (VM instances) in cloud computing environments with the objective of maximizing the revenue of the service provider as well as providing an efficient allocation of resources. A recent survey on market-oriented resource management and scheduling in computing services can be found in Garg and Buyya (2011).

5 System Design Issues and Research Opportunities

By adopting the agent-based approach, the challenges of distributed environment and complicated multiple objectives in service scheduling have been naturally modelled in the agent-oriented architecture. The main design issue is how to design agent-based scheduling systems such that they can effectively address the challenges of dynamic scheduling environment and the presence of customers' private information. In the previous section, we have reviewed typical agent-based scheduling approaches aiming at addressing these challenges from a domain specific perspective. In this section, we summarize the existing agent-based service scheduling approaches from the system design perspective and identify future research opportunities

5.1 System structures

Existing literature on agent-based service scheduling system design usually adopt the physical decomposition approach for agent encapsulation. Service providers who control the service resources are modeled as provider agent. Users who request services are modeled as customer agents. In some cases, such as carrier collaboration in transportation services, a service provider can also request services from other providers. In this situation, a service provider can have both the roles of provider agent and customer agent. Given the agent encapsulation scheme, agent system architectures provide the organizing framework within which agents interact with each other. In the context of agent-based service scheduling, two types of system structures are usually adopted, namely mediated structure and autonomous structure. Mediated structure utilizes a mediator to coordinate the allocation of resources to users. Service provider agent often assumes the role of mediator. For example, in healthcare scheduling, provider (resource) agents usually take the role of mediator and coordinate the resource allocation among patients (Paulussen et al., 2003; Hannebauer and Muller, 2001; Hosseini et al., 2011).

Autonomous structure appears in the settings where a service provider also requires services from other providers, that is, an agent is both a provider and a customer. In autonomous structure, interactions between agents are not coordinated by mediator agents. Instead, agents optimize their schedules by exchanging their resources (Krajewska and Kopfer, 2006b, Gujo et al., 2009). In some service scheduling settings, such as meeting scheduling or workforce scheduling, there are no explicit resource times to be allocated. Instead, the main issue is to find a meeting time or work schedule which is agreeable by all participants. For example, in Becker and Hans (2006), agents representing operation room staffs negotiate with each other based on the Nash bargaining solution to schedule their work shifts. Autonomous structure is also often used in agent-based meeting scheduling applications (Hassine et al., 2004, Modi et al., 2004, and Franzin et al., 2002).

5.2 Negotiation mechanisms

Given its inherently decentralized nature, agent-based service scheduling must coordinate agents' behavior using some types of negotiation protocols. Among others, the Contract Net protocol (CNP) and economic based models, such as auctions, are more prevalent. CNP is essentially a general tendering procedure. However, unlike auctions, the awarding decision may not be related to price or cost factors. To summarize, each agent (manager) having work to subcontract broadcasts a call for bidding message and waits for other agents (contractors) to send back their bids. After receiving bids from all agents or waiting for a certain time period, the manager evaluates all bids received based on its evaluation criteria and awards its contracts to one or more contractors, which then process the subtask. CNP coordinates task allocation, providing dynamic allocation and natural load balancing. Unlike general equilibrium market mechanisms or auctions, which usually require a mediator, contract nets are purely distributed model, in which any agent can act as a manager and subcontract tasks to other agents. CNP can be easily embedded into the autonomous system structure and is suitable for distributed dynamic scheduling. For example, in Zhiming (2011), CNP is used to dynamically reallocate tasks among agents in an operation rooms scheduling setting. The drawback of CNP is that there is no built in mechanism to motivate agents to reveal their private information. Therefore, it is not sufficient in the service scheduling settings where there is the presence of customers' private information.

Auctions can accommodate customer private information by providing necessary incentives to customers. There is a wealth of literature on auction design. Different auction formats such as sequential auctions, simultaneous auctions and combinatorial auctions have been studied extensively in the literature. In agent-based service scheduling, combinatorial auctions (also called bundle auctions) are usually used because scheduling is, in its essence, a combinatorial optimization problem. Typical examples include various implementations of VCG auctions (Crawford & Veloso, 2004; Sheffi, 2004; Berger and Bierwirth, 2010). However, due to high computational complexity, VCG is not practical for large scale problems, especially in dynamic environments. To provide better responsiveness sequential auctions, simultaneous auctions and iterative implementations of combinatorial auctions are also adopted in services scheduling (Paulussen et al., 2003; Song and Regan, 2003; Sönmez & Ünver, 2007; Kwon et al., 2005; Gujo et al., 2009). We will compare different auction models and analyze their applicability to agent-based service scheduling in the following subsection.

5.3 Research opportunities

This paper provides a survey on system design for service process scheduling. Our review covers several representative service domains. The reviewed approaches focus on either dynamic scheduling environment or users' private information. These approaches may not be sufficient for many real world service scheduling applications because they usually deal with only part of the challenges. Based on this survey, as well as on our first-hand research and development experience in this area, we believe that future research on an integrated approach that tackles service scheduling challenges concurrently is much needed. While there is no built in mechanism in CNP to address customers' private information, a logical step to the integrated approach is to design auctions which can accommodate dynamic changes and handle bundles of resource requirements in service scheduling. The key issue is how to deal with enormous computational complexities of combinatorial auctions in dynamic environments.

In general auction terms, combinatorial auctions (CA) allow bidders to place bids on bundles of items. It addresses bundle preferences explicitly. However, the computation required to solve hard valuation problems and winner determination problems can be prohibitive. In general, CAs are likely to be practical for smaller size problems. In addition, CAs require a complete valuation on alternative schedules to be revealed to the auctioneer. In service scheduling, customers are often reluctant to do so in case information might leak out and adversely affect their other decisions or negotiations. Lack of transparency is another practical concern in CAs. It can be difficult to explain to the customers why a certain schedule is chosen. Iterative bundle auctions are iterative implementations of CAs. This class of auction has practical significance because it addresses the computational and informational complexities of CAs by allowing bidders to reveal their preference information only as necessary as the auction proceeds, and bidders are not required to submit (and compute) complete and exact information about their private valuations. In many cases, iterative auctions present better computational and privacy properties than those of CAs. In addition, iterative auctions have the potential of accommodating dynamic events, which is an important requirement in service scheduling applications. With a careful design of the structure and components, iterative bundle auctions have the potential of significantly reducing computational costs and accommodating the dynamic environment and users' private information in service scheduling.

Differently from CAs and their iterative implementations, sequential and simultaneous auctions price bundles as the sum price of the individual items. However, they do not allow bidders to bid on bundles of items. Sequential auctions suppose that the set of items is auctioned in sequence. Bidders bid for items in a specific known order and can choose how much (and whether) to bid for an item depending on past successes, failures, prices and so on. Sequential auctions are particularly useful in situations where setting up combinatorial or simultaneous auctions is infeasible. Simultaneous auctions sell multiple items in separate markets simultaneously. Bidders have to interact with simultaneous but distinct markets in order to obtain a combination of items sufficient to accomplish their task. Real-world markets quite typically operate separately and concurrently despite significant interactions in preferences. Sequential and simultaneous auctions tackle the complementarities over resources in the same spirit of general equilibrium theory. These auctions fail when there are no prices that support an efficient solution (the existence problem) and also when agents bid cautiously to avoid purchasing an incomplete bundle (the exposure problem). However, given that these auctions are more practical in terms of computation, they are two important models worthy of further study.

In addition to the design of core negotiation mechanisms, there are other research needs in agent-based service scheduling. For example, there is a lack of systematic analysis and comparison on how system design factors affect computational time in agent-based service scheduling systems. To adequately test and evaluate various approaches, benchmark problems are also needed. Furthermore, the systems must be designed to integrate a wide range of real-time information and uncertain parameters into the dynamic service scheduling process. Differently from existing auction designs in the literature, dynamic pricing cannot be applied to some services, such as healthcare and government services. In these settings, bidding based service scheduling systems without dynamic pricing are needed. We believe this is an interesting research topic even for auction design in general.

6 Conclusion

Service scheduling are inherently distributed and dynamic. The presence of customers' private information imposes additional challenges in finding high quality solutions. Agent-based systems can be an appropriate approach to service scheduling due to their distributed and autonomous nature. This paper analyzed challenges in service scheduling system design and reviewed agent-based scheduling approaches in representative service domains through the lenses of how they address the challenges of service scheduling. Despite of many domain specific design applications in agent-based service scheduling, there is a lack of general problem formulations, classifications, solution frameworks, and test beds. Constructing these general models for service scheduling will greatly facilitate the collaboration of researchers in this area and guide the effective development of integrated service scheduling systems. Moreover, the applicability of a service scheduling approach to industrial settings will largely depend on how it copes with distributed and dynamic environments and on how it computes high quality solutions despite the presence of customers' private information.

References

Abdennadher, S., & Schlenker, H. (1999). INTERDIP-an interactive constraint based nurse scheduler. *Proceedings of the Eleventh Conference on Innovative Applications of Artificial Intelligence, Menlo Park, CA,* 838-843

Aickelin, U., & Dowsland, K. A. (2001). Exploiting problem structure in a genetic algorithm approach to a nurse rostering problem. *Journal of Scheduling 3(3)*, 139-153

An, B., Lesser, V., Irwin, D., & Zink, M. (2010). Automated negotiation with decommitment for dynamic resource allocation in cloud computing. *Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems*, 981–988

Azaiez, M. N., & Sharif, S. (2005). A 0-1 goal programming model for nurses cheduling. *Computers and Operations Research 32(3)*, 491-507

Bailey, R. N., Garner, K. M., & Hobbs, M. F. (1997). Using Simulated Annealing and Genetic Algorithms to Solve Staff Scheduling Problems. *Asia-Pacific Journal of Operational Research 14(2)*, 27-43

Bannock, G., Baxter, R. E., & Reese, R. (1982). *The Penguin Dictionary of Economics*. Penguin Books, Ltd., Harmondsworth, Middlesex England

Becker, M., & Hans, C. (2006). Artificial Software Agents as Representatives of Their Human Principals in Operating-Room-Team-Forming. *Multi-agnet Engineering International Handbooks on Information Systems,* 221-237

Berger, S., & Bierwirth, C. (2010). Solutions to the request reassignment problem in collaborative carrier networks. *Transportation research Part E,Volume 46,No.5*, 627-638

Burke, E. K., Elliman, D. G., & Weare, R. F. (1995). A hybrid genetic algorithm for highly constrained timetabling problems. *Proceedings of the 6th International Conference on Genetic Algorithms, Pittsburgh, USA,Morgan Kaufmann, Los Altos, CA,* 605-610

Burke, E. K., & Newall, J. P. (1999). A multi-stage evolutionary algorithm for the timetable problem. *IEEE Transactions on Evolutionary Computation 3 (1),* 63–74

Burke, E. K., Newall, J. P., & Weare, R. F. (1996). A memetic algorithm for University exam timetabling. *Burke and Ross*, 241-250

Burke, E. K., Newall, J. P., & Weare, R. F. (1998). Initialisation strategies and diversity in evolutionary timetabling. *Evolutionary Computation 6 (1)*, 81-103 (special issue on Scheduling)

Crawford, E., & Veloso, M. (2004). Mechanism Design for Multi-Agent Meeting Scheduling Including Time Preferences, Availability, and Value of Presence. *Proceedings of the IEEE/WIC/ACM International Conference on Intelligent Agent Technology (IAT)*

Davidsson, P., Henesey, L., Ramstedt, L., T¨ornquist, J., & Wernstedt, F. (2005). An analysis of agent-based approaches to transport logistics. *Transportation Research, Part C, 13,* 255–271

Davis, L. (1985). Job shop scheduling with genetic algorithms. *Proc. 1st int. Conf. on Genetic algorithms and their Applications, Pittsburgh, PA,* 130-140

Demeester, P., Souffriau, W., De Causmaecker, P., & Vanden Berghe, G. (2010). A hybrid tabu search algorithm for automatically assigning patients to beds. *Artif. Intell. Med. 48(1)*, 61–70

Dowsland, K. (1998). Nurse scheduling with tabu search andstrategic oscillation. *European Journal of Operational Research 106 (2–3),* 393–407

Fischer, K., Müller J. P. and Pischel, M. (1995).Cooperative transportation scheduling: an application domain for DAI. *Journal of Applied Artificial Intelligence*

Franzin, M. S., Freuder, E. C., & Rossi, F. (2002). Multi-agent meeting scheduling with preferences: efficiency, privacy loss, and solution quality. *American Association for Artificial Intelligence AAAI*

Gagliano, R. A., Fraser, M. D., & Schaefer, M. E. (1995). Auction allocation of computing resources. *Communications of the ACM, 38 (6)*, 88–102

Garg, S., and Buyya, R. (2011). Market-Oriented Resource Management and Scheduling: A Taxonomy and Survey, *Cooperative Networking 277-306, M. S. Obaidat and S. Misra (eds),* ISBN: 978-0-470-74915-9, Wiley Press, New York, USA

Ghaemi, M.,Vakili, M., & Aghagolzadeh, A. (2007). Using a genetic algorithm optimizer tool to solve university timetable scheduling problem. *9th international symposium on signal processing and its Application*

Gomber, P., Schmidt, C., Weinhardt, C. (1997). Elektronische Märkte für die dezentrale Transportplanung, *Wirschaftsinformatik 39(2),*137-145

Grano, M., Medeiros, D. J., & Eitel, D. (2009). Accommodating individual preferences in nurse scheduling via auctions and optimization. *Health Care Manage Science, Volume 12,*228-242

Groothuis, S., & Merode, G. (2001). Simulation as decision tool for capacity planning. *Journal of Computer Methods and Programs in Biomedicine 66* , 139–151

Gueret, C., Jussien, N., Boizumault, P., & Prins, C. (1995). Building University Timetables Using Constraint Logic Programming. *Proc. of the 1st Int. Conf. on the Practice and Theory of Automated Timetabling,* 393- 408

Gujo, O., Schwind, M., & Vykoukal, J. (2009). A combinatorial intra-enterprise exchange for logistics services. *Information systems and e-business management,Volume 7,No 4,*447-471

Gunawan, A., Ming, K., & Poh, K. (2007). Solving the teacher assignment-course scheduling problem by hybrid Algorithm. *International journal of Computer, information and system science and engineering, 1(2),*139-141

Hancock, W.M., & Walter, P. F. (1984). The use of admissions simulation to stabilize ancillary workloads. *Simulation journals,* 88-94

Hannebauer, M., & Muller, S. (2001). Distributed Constraint Optimization for Medical Appointment Scheduling. *Proceedings of the fifth international conference on autonomous agents,*139 -140

Harvey, J. (1998). Service quality: A tutotial. *Journal of Operations Management 16(5)*, 583-597

Hassine, A. B., Defago, X., & Ho, T. B. (2004). Agent-Based Approach to Dynamic Meeting Scheduling Problems. *Proceedings of the Third International Joint Conference on Autonomous Agents and Multiagent Systems,Volume 3,* 1132 -1139

Henz M., & Wurtz J. (1995). Using Oz for college timetabling. *Proceedings of the 1st Int. Conference on the Practice and Theory of Automated Timetabling,* 283- 296

Hertz, A. (1991). Tabu Search for Large Scale Timetabling Problems. *European Journal of Operational Research 54,* 39-47

Hertz, A. (1992). Finding a Feasible Course Schedule Using Tabu Search. *Discrete Applied Mathematics 35(3),* 255-270

Ho, Ch., & Lau, H. (1999). Evaluating the impact of operating conditions on the performance of appointment scheduling rules in service systems. *European Journal of Operational Research 112* ,542-553

Hosseini, H., Hoey, J., & Cohen, R. (2011). Multi-Agent Patient Scheduling Through Auctioned Decentralized MDPs. *Proceedings of the 6th InformsWorkshop on Data Mining and Health Informatics*

Hur, D., Mabert, V. A., & Bretthauer, K. M.(2004). Real-time work schedule adjustment decisions: An investigation and evaluation. *Production and Operations Management 13(4),* 322

Jack, E. P., & Powers, T. L. (2004). Volume flexible strategies in health services: A research framework. *Production and Operations Management 13(3),* 230

Jaumard, B., Semet, F., & Vovor, T. (1998). A generalized linear programming model for nurse scheduling. *European Journal of Operational Research 107(1),*1-18

Jennings, N. R. (2001). An agent-based approach for building complex software systems. *Communications of the ACM, 44(4),*35- 41

Kirkpatrick, S., Gelatt, C. D., & Vecchi, M. P. (1983). Optimization by Simulated Annealing. *American Association for the Advancement of Science New Series, Vol. 220, No. 4598.,* 671-680

Kotler, P., & Keller, K. (2006). *Marketing management, Twelfth edition.* Prentice-Hall, Upper Saddle River, New Jersey

Krajewska, M. A., & Kopfer, H. (2006a). Profit sharing approaches for freight forwarders: An overview", *Proceedings of the 5th International Conference on Modern Trends in Logistics,*157-161

Krajewska, M. A., & Kopfer, H. (2006b). Collaborating freight forwarding enterprises, request allocation and profit sharing. *OR spectrum, Volume 28, No2,* 301-317

Krishna, A., & Ünver, M. U. (2007). Improving the Efficiency of Course Bidding at Business Schools: An Experimental Study. *Marketing Science, forthcoming*

Kwon, R. H., Lee, C., & Ma, Z. (2005). An integrated combinatorial auction mechanism for truckload transportation procurement. *Technical Report, Mechanical and Industrial Engineering, the University of Toronto, Ontario, Canada*

Lang, N., Moonen, H. M., Srour, F. J., & Zuidwijk, R. A. (2008). Multi Agent Systems in Logistics: A Literature and State-of-the art Review. *ERIM Report Series,* Reference No. ERS-2008-043-LIS

Meisels, A., & Kaplansky, E. (2003). Scheduling Agents – Distributed Timetabling Problems. *Lecture Notes in Computer Science, Practice and Theory of automated timetabling IV,Volume 2740/2003,* 166-177

Modi, P.,Veloso, M., Smith, S. F., & Oh, J. (2004). CMRadar: A Personal Assistant Agent for Calendar Management. *Lecture Notes in Computer Science, LNCS 3508,*169–181

Paechter, B., Cumming, A., & Luchian, H., (1995). The use of local search suggestion lists for improving the solution of timetabling problems with evolutionary algorithms. *Proceedings of the AISB Workshop on Evolutionary Computing, Sheffield, England.*

Paechter, B., Cumming, A., Norman, M. G., & Luchian, H. (1996). Extensions to a memetic timetabling system. *The Practice and Theory of Automated Timetabling, volume 1153 of Lecture Notes in Computer Science. Springer Verlag,* 251–265

Paulussen, T. O., Jennings, N. R., Decker, K. S., & Heinzl, A. (2003). Distributed patient scheduling in Hospital. *Coordination and Agent Technology in Value Networks, GITO*

Pearce, D. W. (1981). *The dictionary of modern economics.* The MIT Press, Cambridge, Massachusetts

Petrovic, D., Morshed, M., & Petrovic, S. (2011). Multi-objective genetic algorithms for scheduling of radiotherapy treatments for categorized cancer patients. *Journal of Expert Systems with Applications,38(6), 6994-7002*

Pinedo, M.L (2009). *Planning and scheduling in manufacturing and services* (2nd ed.). Springer, New York. doi: 10.1007/978-1-4419-0910-7

Sampson, S. E. & Froehle, C. M. (2006). Foundations and implications of a proposed unified services theory. *Production and Operations Management,* 329-343

Sampson, S. E. (2001). *Understanding service businesses: Applying principles of the unified services theory* (2nd ed.). John Wiley & Sons, New York, New York

Schönberger, J. (2005). *Operational Freight Carrieer Planning*. Springer, Berlin

Schönsleben P., Hieber R. (2004). Gestaltung von effizienten Wertschöpfungspartnerschaften im Supply Chain Management. *Busch A., Dangelmaier W., Integriertes Supply Chain Management, Wiesbaden.*

Sheffi, Y. (2004). Combinatorial Auctions in the Procurement of Transportation services. *Interfaces,Volume.34* , 245-252

Shen, W., Wang, L., & Hao, Q. (2006). Agent-based distributed manufacturing process planning and scheduling : a state-of-the-art survey. *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews 36(4),* 563-577

Sim, K. M. (2012). Complex and Concurrent Negotiations for Multiple Interrelated e-Markets. *IEEE Transactions on Systems, Man and Cybernetics, Part B: Cybernetics, PP(99),* doi: 10.1109/TSMCB.2012.2204742, 1-16

Singh, A., & Malhotra, M. (2012). Agent Based Framework for Scalability in Cloud Computing. *International Journal of Computer Science & Engineering Technology (IJCSET), 3(4),* 41-45

Song, J., & Regan, A. C. (2003). An Auction Based Collaborative Carrier Network.*Technical report: UCI-ITS-WP-03-6, Institute of Transportation Studies, University of California, Irvine*

Sönmez, T., & Ünver, U. (2007). *Course Bidding at Business Schools.* Retrieved from http://ssrn.com/abstract=1079525 2007

Wainer, J., Ferreira, P., & Constantino, E. R. (2007). Scheduling meetings through multi-agent negotiations. *Decision Support Systems 44,* 285–297

Wall, B. M. (1996). *A Genetic Algorithm for Resource-Constrained Scheduling,* Ph.D. thesis, Massachusetts institute of technology

Wang, C. (2007). Economic Models for Decentralized Scheduling. Ph.D thesis. University of Western Ontario.

Wang, W., & Gupta, D. (2011). Adaptive Appointment Systems with Patient Preferences. *Manufacturing and Service Operations Management 13(3),* 373-389

Wemmerlov, U. (1990). A taxonomy for service processes and its implications for system design. *International Journal of Service Industry Management 1(3),* 13–27

Wolski, R., Plank, J. S., Brevik, J., & Bryan, T. (2001). Analyzing market-based resource allocation strategies for the computational grid. *International Journal of High Performance Computing Applications, 15 (3),* 258-281

Zaman, S., & Grosu, D. (2011). Combinatorial Auction-Based Dynamic VM Provisioning and Allocation in Clouds. *IEEE Third International Conference on Cloud Computing Technology and Science (CloudCom)*,107-114

Zhiming, Z. (2011). A Two-stage Scheduling Approach of Operation Rooms Considering Uncertain Operation Time. *International Conference on Information Science and Technology,Nanjing, Jiangsu, China 250.* doi: 10.1115/DETC2011-48263

Author Biographies

Farnaz Dargahi is a Ph.D candidate at The Department of Electrical and Electronics Engineering, Concordia University, Montreal, Quebec, Canada. Her research focuses on economic-based models for service systems optimization.

Chun Wang is an assistant professor at Concordia Institute for Information Systems Engineering, Concordia University, Montreal, Quebec, Canada. His research focuses on distributed scheduling, multi-agent systems, e-Supply Chain, e-Commerce and algorithmic mechanism design.

Mohammad F. H. Bhuiyan is a Ph.D. student at Concordia Institute for Information Systems Engineering, Concordia University, Montreal, Quebec, Canada. His research interests include distributed systems, large scale system optimization, mechanism design and market-based models for computing services management.

Hamidreza Mehrizi is a Master’s student at Concordia Institute for Information Systems Engineering, Concordia University, Montreal, Quebec, Canada. His research focuses on auction-based models for carrier collaboration in transportation services.