Online Advance Scheduling with Patient Preferences

XINSHANG WANG - VAN-ANH TRUONG - DAVID BANK
COLUMBIA UNIVERSITY
JUNE 10, 2015
How does a patient book a medical appointment online today?
Schedule an Appointment

Step 1: Choose reason for visit

Currently this functionality is only available for appointments with physicians in Weill Cornell Medical Associates. To request an appointment with another physician, please use the Request Appointment feature from the Appointments menu. We hope to make this functionality available for all providers in the near future.

Please choose a reason for visit and click Continue. If you would like to schedule an appointment for a different reason, please call the practice.

Choosing a visit type:
- **Non Urgent Visit**: this is a routine followup appointment for a chronic medical condition or new, non-urgent medical problem (generally schedules in 5 days or longer).
- **Well Visit/Physical**: this is your annual well visit/physical exam, routine annual gyn exam, or well-child visit for routine vaccines. Some insurances allow different intervals between routine physicals exams and usually must be at least 365 days from the last one so please confirm with your insurance company if you are unsure.
- **Flu vaccine**: this is a nurse-only visit to receive your seasonal flu vaccine only.

FOR ALL OTHER VISIT TYPES (URGENT SICK VISIT, PROCEDURES, NURSE-ONLY VISITS FOR BLOOD TESTS OR NON-FLU VACCINES) PLEASE CALL THE OFFICE.

Reason for visit:

- Non Urgent Visit (not annual check-up)
- Non Urgent Visit (not annual check-up)
- Well Visit/Physical
- Flu vaccine

[Continue] [Cancel]
Schedule an Appointment

Step 3 of 6: Choose provider

Reason for visit: Non Urgent Visit (not annual check-up)


Please note that you must arrive 10 minutes before your scheduled appointment time to update your insurance and demographic information and complete any necessary forms. If you feel you need a more urgent appointment please call the office.

Please choose the provider with whom you want to schedule an appointment and click Continue.

Schedule at: Internal Medicine

Schedule with: Chang, Joseph J, MD - PCP - (Weill Cornell Medical Associates)


Preferred times:

<table>
<thead>
<tr>
<th></th>
<th>Monday</th>
<th>Tuesday</th>
<th>Wednesday</th>
<th>Thursday</th>
<th>Friday</th>
</tr>
</thead>
<tbody>
<tr>
<td>Morning</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td></td>
</tr>
<tr>
<td>Afternoon</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✔️</td>
</tr>
</tbody>
</table>
Schedule an Appointment

Step 4 of 6: Choose appointment time slot

Reason for visit: Non Urgent Visit (not annual check-up)

Please note that you must arrive 10 minutes before your scheduled appointment time to update your insurance and demographic information and complete any necessary forms. If you feel you need a more urgent appointment please call the office.

This page displays the date and time of the available time slots for the selected provider. Please choose the desired time slot and click Continue > to view the complete appointment information before scheduling an appointment.

If you would like to see more time slots, click Next to continue to the next set of available time slots, or click < Back and change the date and time range.

<table>
<thead>
<tr>
<th>Date/Time</th>
<th>Providers/Resources</th>
<th>Department</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tuesday October 07, 2014 11:20 AM</td>
<td>Chang, Joseph J, MD</td>
<td>Weill Cornell Medical Associates</td>
<td>Internal Medicine</td>
</tr>
</tbody>
</table>

Previous Times  Next Times

< Back  Continue >  Cancel
Schedule an Appointment (Anjana)

Please note that you must arrive 10 minutes before your scheduled appointment time to update your insurance and demographic information and complete any necessary forms. If you feel you need a more urgent appointment please call the office.

Weill Cornell CONNECT did not find any available time slots which meet your requirements.

Click < Back to change your current scheduling preferences. Click Cancel to start over from the beginning.

There are no time slots available.

< Back  Cancel
MyChart

- An e-scheduling system developed by Epic Systems
  - an electronic medical records company
  - managing the records for millions of healthcare providers: 69% of Stage 7 U.S. Hospitals, 71% of children's hospitals, and 83% of Stage 7 Clinics
  - Reaches 54% of U.S. patient population

Growth of e-scheduling

- MyChart is one of many e-scheduling systems
- Increasing trend for reservation of services over internet portals
  - 60% percent of people consider digital services such as online appointment scheduling are an important criterion in selecting a care provider.
  - 50% of 18-24 year-old respondents wanted mobile access for scheduling appointments

Source: Google Consumer Insights Survey 2015
Growth of e-scheduling

- Another example, OpenTable
  - A website that matches diners with restaurants and enables direct reservation
  - Used by 15 M diners across more than 31,000 restaurants each month
E-scheduling challenges

1. Random, non-stationary arrivals

Weekly arrivals in Clinical Genetics, CUMC
E-scheduling challenges

2. Differing priorities

- Eg: Urgent versus non-urgent in primary care
- Urgent patients might need to call in rather than book online
- System needs to operate seamlessly across different modes of access
2. Differing availability and preferences
   - Time of day
   - Time of week
   - Location
   - Primary-care provider versus other providers
   - Ability to learn and accommodate preferences determine usability and eventual adoption
2. Expiration of unused appointments slots

- Unused slots expire at the end of a day
- Want to fill patients into soon-to-expire slots
- But also want to accommodate patient preferences
5. **No-shows and cancellations**

Show probability in Clinical Genetics, CUMC
Past works

- Gupta and Wang 2008
  - Focuses on a single day
  - Distribution of preferences for time and provider
  - Each patient makes a single request for a slot
  - Clinic accepts or rejects request. No recourse.
  - DP intractable. Heuristics
Past works

- Feldman, Liu, Topaloglu and Ziya (2015)
  - Multi-period, stationary setting
  - 1 patient type, 1 provider, no-shows
  - Preferences for slots follows MNL model
  - Provider displays slots as patients arrive. Patients choose.
  - Intractable. Structural results and heuristics
Our new model

- Multi-period, **non-stationary** setting
- Multiple providers, locations, patient types, etc.
- Arbitrarily complex patient preferences
- Slot-dependent no-shows
Our theoretical performance characterization

- First algorithms with theoretical guarantee on performance
- Algorithms are asymptotically optimal
- Excellent empirical performance as tested on real data
Our new general algorithms and bounds

- Connections to online matching
- New algorithms and bounds for online matching
- Application in revenue management, management of opaque products, display-ad allocation, general resource allocation
Model

- Continuous-time, finite time horizon

- $n$ known appointment slots over the horizon
  - Each is associated with a time, provider, location, etc.

- $m$ patient types

- Type $i$ arrival process is non-homogenous Poisson with rate $\lambda_i(t)$
Model

- Type of a patient observed upon arrival

  - Recall MyChart example

  - Uses patient profile and input at time of booking

  - Similar to booking with human agents
**Model**

- \( r_{ij} \): benefit value of assigning patient of type \( i \) to appointment slot \( j \)

- Captures patient preference, hospital revenue, priority, etc.

- Patients arriving at \( t \) have 0 preference for slots earlier than \( t \)

  - Models perishability of slots
Model

- One available slot is assigned, or the patient is rejected or diverted

- Rejection inevitable when demand exceeds capacity

- Or when capacity must be reserved for higher priority patients who might come later
Our model

Appointment slots

patient

patient
Our goals

- Maximize total benefit of an allocation
- Computing an optimal policy is intractable
- We aim to design a scheduling policy that achieves at least $\frac{1}{2}$ of the expected objective value of OFF

  - OFF is an optimal offline policy. Knows all information upfront and makes optimal decisions
Connections with online matching literature

- No non-trivial bound for adversarial demand

- Analysis of arrivals drawn randomly with or without replacement from some fixed set
  
  Jaillet and Lu (2014), Manshadi et al. (2012), Bahmani and Kapralov (2010), Feldman et al. (2009), Haeupler et al. (2011)
  
  Kleinberg (2005), Babaio et al. (2008), Goel and Mehta (2008), Mahdian and Yan (2011), Karande et al. (2011).

- Analysis with small-bid assumption
  

- We offer a new bound and analysis for problems with non-stationary arrivals
Our policy

- Step 1: Solve a deterministic assignment LP (DLP)

- $x_{ij}$: average number of type $i$ patients to be assigned appointment slot $j$
Our policy

- Let $x^*$ be an optimal solution to the assignment LP

- Step 2: Stream arriving patients of type $i$ to slot $j$ with probability

\[ \frac{x_{ij}^*}{\Lambda_i} \]
Our policy

- This streaming separates problem into slot-by-slot admission-control problems

- There is a demand stream coming to each slot

- Demand of type-i patients for slot j is Poisson with rate $\lambda_{ij}(t) = \lambda_i(t)x_{ij}^*/\Lambda_i$

- (Separated) problem for the slot is whether to accept/reject each in coming patient
Our scheduling policy

- Step 3: Solve the separated problems using dynamic programing

- HJB equation that determines the optimal future benefit $f_j(t)$ for slot $j$ given streamed arrival rates

$$f'_j(t) = -\sum_{i=1}^{m} \lambda_i(t) x^*_i / \Lambda_i \cdot (r_{ij} - f_j(t))^+,$$

Fraction of type $i$ patients who are assigned appointment slot $j$
Proof of performance guarantee

- Easy to shown that DLP gives an upper bound on expected revenue of OFF

\[
\begin{align*}
\text{max} & \quad \sum_{i=1}^{m} \sum_{j=1}^{n} x_{ij} r_{ij} \\
\text{s.t.} & \quad \sum_{j=1}^{n} x_{ij} \leq \Lambda_i, \quad \text{for } i = 1, 2, \ldots, m \\
& \quad \sum_{i=1}^{m} x_{ij} \leq 1, \quad \text{for } j = 1, 2, \ldots, n \\
& \quad x_{ij} \geq 0.
\end{align*}
\]

- OFF gets at most \( \sum_{ij} x_{ij}^* r_{ij} \)
Proof of performance guarantee

- Focus on a resource $j$
- OFF gets at most $\sum_i x_{ij}^* r_{ij} = \sum_i \int_0^1 \lambda_{ij}(t) r_{ij}$
- Show that ON gets at least half as much
- Technique: \textit{dynamics constrained global bounding}
Proof of performance guarantee

\[
\inf_{\lambda_{ij}(t), r_{ij}, i=1,...,m} f_j(0)
\]

\[
\text{s.t. } f_j'(t) = -\sum_{i=1}^{m} \lambda_{ij}(t) \cdot (r_{ij} - f_j(t))^+
\]

\[
\int_0^1 \sum_{i=1}^{m} \lambda_{ij}(t) dt \leq 1
\]

\[
\int_0^1 \sum_{i=1}^{m} \lambda_{ij}(t) r_{ij} dt = 1
\]

\[
\lambda_{ij}(t), r_{ij} \geq 0. \ i = 1,2,...,m
\]

\[
f(1) = 0,
\]

Dynamic program

Capacity constraint for slot j from DLP

Upper bound on OFF's Revenue, normalized to 1

Boundary condition
Proof of performance guarantee

- A lower-bound on the optimization problem is 0.5
- Thus competitive ratio is at least 0.5 for slot j
- Repeat for all slots to get overall competitive ratio of 0.5
Potential problem

- Random streaming of arriving customers might create **unfair** situations

  - High-priority patient gets rejected while low-priority patient gets accepted

  - Even when preferences of patients are similar
Modification to the policy

- Solve for the benefit functions $f$ as before

- Step 4: When a patient of type $i$ arrives at time $t$, assign him to appointment slot $j$ such that
  \[ j = \max_{k} \{ r_{ik} - f_k(t) : r_{ik} \geq f_k(t) \} \]

- Reject the patient if the set is empty

- Can show that competitive ratio still holds
Can we do better than 0.5?

- A simple example shows that 0.5 is the best possible ratio
  - One resource, 2 customer types

- Thus, ON is an optimal online algorithm
Asymptotic performance

- Can show that **ON is asymptotically optimal** when the system size grows large

\[ \lim_{\theta \to \infty} \frac{\sum_{j=1}^{n} f_j(0, C_j)}{r^\tau x^*} = \lim_{\theta \to \infty} \frac{\sum_{j=1}^{n} f_j(0, \theta \tilde{C}_j)}{\theta r^\tau \tilde{x}^*} = 1, \]

where \( x^* \) is an optimal solution to (24).
No-shows and overbooking

- Assume each customer has no-show probability of $p_j$ when assigned to slot $j$ and overbooking cost $D_j$ when denied slot $j$

- Treat each overbooked slot as a virtual slot

- Benefit in assigning a virtual slot is smaller due to expected overbooking cost

- Same algorithm and competitive ratio
Computing the policy in practice

- LP potentially too large to solve directly when there are many patient types
- Can calculate “value functions” by simulation
- Idea: sample a subset $S$ of customer types and solve a small LP corresponding to these types
- Use the dual to construct a primal solution to original LP
Computing the policy in practice

Let $\epsilon = |S|/m$

**Theorem 7.** (Feldman et al. 2010) With high probability,

$$\sum_{i=1}^{m} \sum_{j=1}^{n} r_{ij} x_{ij}(p) \geq (1 - O(\epsilon)) \sum_{i=1}^{m} \sum_{j=1}^{n} r_{ij} x_{ij}^*, $$

given

$$\max_{i,j} \left\{ \frac{r_{ij} \Lambda_i}{\sum_{kl} r_{kl} x_{kl}^*} \right\} \leq \frac{\epsilon}{(n+1)(\ln m + \ln n)}$$

and

$$\max_{i} \{ \Lambda_i \} \leq \frac{\epsilon^3}{(n+1)(\ln m + \ln n)}.$$
Computing the policy in practice

- Let $\epsilon = |S|/m$

**Theorem 8.** If $x(p)$ is $1 - O(\epsilon)$ optimal for the LP (2), then if we use $\mathbb{E}[\hat{f}_j(t)|S]$ as the derivative of benefit function, our algorithms are $0.5(1 - O(\epsilon))$-competitive.
Numerical studies

- 12 weeks of appointment scheduling data within a 5-year data set
- Collected from Division of Clinical Genetics, Columbia University Medical Center
- Each entry in the data records:
  - the date that patient arrives to make an appointment
  - the exact time of the appointment
  - whether the patient eventually showed up to the original appointment, canceled the appointment some time later, or missed the appointment
Numerical studies

Probability that the patient arriving in period $i$ will show up in slot $r_{ij} = j$ without canceling the appointment some time later or missing the appointment eventually.

- Reward of an assignment == show probability
- Show probabilities depend on
  1. Appointment lag
  2. Time of day (morning/afternoon)
  3. Time of week
Numerical studies

- Baseline policies
  1. Greedy policy
  2. Bid-price heuristic
  3. Actual strategy

- Actual policy reuses cancelled slots. We do not.
  - Even with this disadvantage we still do well
Numerical studies

Experiment 1: Patients identical except for time of arrival

<table>
<thead>
<tr>
<th>Scheduling Policy</th>
<th>Percentage of patients who show up to original appointment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actual Strategy</td>
<td>58%</td>
</tr>
<tr>
<td>Greedy</td>
<td>70%</td>
</tr>
<tr>
<td>Bid-Price Heuristic</td>
<td>77%</td>
</tr>
<tr>
<td>Marginal Allocation Algorithm</td>
<td>79%</td>
</tr>
</tbody>
</table>
**Numerical studies**

**Experiment 2: With cancellations and overbooking**

Table 2: The total benefit of scheduling policies relative to actual practice under different values of penalty $D$.

<table>
<thead>
<tr>
<th></th>
<th>Greedy</th>
<th>Bid-Price Heuristic</th>
<th>Marginal Allocation Algorithm</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>116.25%</td>
<td>126.90%</td>
<td>132.92%</td>
</tr>
<tr>
<td>3</td>
<td>117.37%</td>
<td>126.25%</td>
<td>134.51%</td>
</tr>
<tr>
<td>4</td>
<td>119.90%</td>
<td>129.78%</td>
<td>138.30%</td>
</tr>
<tr>
<td>5</td>
<td>123.16%</td>
<td>135.04%</td>
<td>143.03%</td>
</tr>
<tr>
<td>6</td>
<td>127.44%</td>
<td>142.64%</td>
<td>147.94%</td>
</tr>
<tr>
<td>7</td>
<td>132.21%</td>
<td>147.40%</td>
<td>154.14%</td>
</tr>
<tr>
<td>8</td>
<td>137.84%</td>
<td>155.06%</td>
<td>161.14%</td>
</tr>
<tr>
<td>9</td>
<td>143.52%</td>
<td>162.22%</td>
<td>168.85%</td>
</tr>
<tr>
<td>10</td>
<td>150.23%</td>
<td>170.72%</td>
<td>176.66%</td>
</tr>
</tbody>
</table>
Numerical studies

Experiment 3: Patients heterogeneous. Simulate availability for each slot uniformly randomly with a fixed probability.

<table>
<thead>
<tr>
<th>$P_A$</th>
<th>Greedy</th>
<th>Bid-Price Heuristic</th>
<th>Marginal Allocation Algorithm</th>
</tr>
</thead>
<tbody>
<tr>
<td>15.00%</td>
<td>100.96%</td>
<td>108.10%</td>
<td>108.65%</td>
</tr>
<tr>
<td>20.00%</td>
<td>110.82%</td>
<td>120.47%</td>
<td>122.50%</td>
</tr>
<tr>
<td>25.00%</td>
<td>115.10%</td>
<td>127.01%</td>
<td>130.01%</td>
</tr>
<tr>
<td>30.00%</td>
<td>116.36%</td>
<td>130.56%</td>
<td>133.86%</td>
</tr>
<tr>
<td>35.00%</td>
<td>117.29%</td>
<td>133.73%</td>
<td>136.84%</td>
</tr>
<tr>
<td>40.00%</td>
<td>118.03%</td>
<td>135.11%</td>
<td>138.75%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Summary

- New model for advance service reservation with consumer preferences
- First online algorithms for this class of problems
- Theoretical performance characterization and good empirical performance
- Implications for online matching and related problems