Integrating the Internet of Things and Big Data Analytics Into Decision Support Models for Healthcare Management

Ralph D. Badinelli
Virginia Tech, Blacksburg, VA, US

Debora Sarno
University of Foggia, Italy
Abstract

This paper presents an analysis of the innovative management of healthcare systems through the application of the Internet of Things (IoT) and Big Data Analytics (BDA). By integrating the technology of IoT and the intelligence of BDA we derive some guidelines for service designs that amplify the personalized, co-creative nature of effective health care. Capturing and interpreting data about patients’ needs and desires, resource availabilities (doctors, nurses, medical equipment, medical supplies and others), treatment options and potential outcomes enables a smart and adaptive health management system for planning, scheduling and coordinating service activities in a jointly managed, co-creative system.

This research formulates the key decision models that will support co-creative service systems that can take advantage of IoT and BDA. We identify the unique tradeoffs that these technologies present to the design and management of healthcare systems, specify the data requirements of the predictive models that are recommended, formulate the mathematical structure of the decision models and describe the changes to the management and culture of the healthcare service system that will be necessary.
1. Introduction

This paper raises some important issues for the implementation of the IoT and BDA in the healthcare industry. Current trends in research are driven by a belief that IoT and BDA engender smart service systems. However, there can be some problems with this belief. IoT and BDA provide a wealth of data to the healthcare service system. Making effective use of these data presents substantial challenges and opportunities. This paper presents these challenges and opportunities and describes the considerations for the design of healthcare systems that ensue.

The healthcare system co-creates value through the coordinated activities of a healthcare supply chain that brings resources to bear on patient needs and the responses of patients to symptoms, treatments, prescriptions and lifestyle choices. From a service-system point of view (Maglio et al., 2009), healthcare systems are configurations of people, information, organizations, and technologies operating together for mutual benefit and common objectives. The traditional approaches to the design of these service system take one of two perspectives:

- a patient-centric view, oriented to patient health (Capunzo et al., 2013) in terms of quality and speed of care;
- a provider view, focused on resource utilization and efficiency (patient waiting time or service level is the only consideration given to the patient - Sarno and Nenni, 2016).

In the new digitization era, the design and management of healthcare services should structurally incorporate both perspectives under the awareness that service itself means value co-creation among the involved actors (Vargo and Lusch, 2004), personalizing the health care experience for each patient and adapting organizational and planning processes to context variability.

Healthcare information systems have always been repositories of a huge amount of data from different sources, ranging from x-ray images to prescription details and much of the data have not been analysed due to the enormous volume and the lack of meaningful correlative analytics (Naqishbandi et al., 2015). In recent years, data has become more and varied through the pervasiveness of Information and Communication Technologies (ICT) and the rapidly increasing use of IoT technology. From Electronic Medical Records to smartphones and wearable devices, a modern healthcare service system will soon have access to a wealth of personal data about each patient. These data include more than just medical data, such as medication regimens, blood pressure, oxygen level, temperature and other diagnostic measurements. Patient location, activity levels, diet, work schedule and other lifestyle information are available to the system. The challenge that is driving a global initiative in healthcare information technology is the integration of these data sources into predictive models of patient prognosis (among others). Much progress has been made in this form of modelling. However, the focus on big data that prevails in current research tends to overlook the ultimate purpose of modelling, which is decision support.

We begin our study of IoT in Healthcare by looking beyond the technology of devices for acquiring data to the value-generating use of the data in supporting healthcare. This perspective is founded on the distinctions among the forms of modelling that comprise a decision support system. Figure 1 illustrates the structural relationships among these model forms. In particular, this figure places predictive analytics in a position of support to the main components (parameters) of the decision model. A consequence of these relationships is a requirement to evaluate the performance of IoT sensors, big data analytics and predictive modelling by the quality of the decisions that can be made from their use.
Large investments in fixed assets and highly trained staffs severely limit the flexibility in capacity of every healthcare service system. The inertia of the supply chain limits its ability to respond to the non-stationary and random nature of demand that is driven by individual patient needs. Therefore, resource allocation methods that can be used in other service industries to respond to variable demand are not as effective in healthcare systems, particularly over short time horizons. The use of IoT and BDA to generate more accurate and dynamic updates of parameters that affect demand and resource availability can enable a different approach known as Demand Response. According to a demand response policy, forecasted demand peaks are managed by service providers by asking customers for shifting or reducing their loads in those periods in exchange for mainly monetary benefits. Electric utilities are a good example of an inertia-constrained supply chain that uses IoT and BDA to enable demand response (Siano and Sarno, 2016) in order to reduce the probability of network congestions and dissipation while saving money that would have been used to buy energy from expensive generators. Although healthcare service systems presently do not adopt the practice of demand response, the necessary technologies are in place and further studies should conceptualize and investigate the types of possible demand response programs.

Indeed, an essential requirement of demand response is also an innovative feature of this practice – demand response requires co-creation since the involved actors choose to exchange services (that in the case of energy can be a disservice in change of money) only for mutual benefit. Through the application of finite capacity scheduling (FCS), a healthcare system can utilize real-time data about patient locations, medical conditions and desires to dynamically assign patients and healthcare resources to medical procedures not just unilaterally and quasi-randomly postponing services but agreeing with patients based on shared healthcare data for improved efficiency in the use of healthcare resources and the personalization of patient care.

In this paper, we review the technologies of IoT and BDA as they can be applied to manage healthcare service systems. Our orientation in evaluating these enabling technologies is that of a service innovator who is concerned only with the role of IoT and BDA in supporting key decisions in the service journey of patients and medical clinics, improving their value proposition in the service-for-service exchange, subsequently
enabling a wiser re-design of healthcare service systems which cannot neglect technology implementations. Specifically, we do not advocate the use of these technologies simply because they are novel, high-tech or exciting. In the final analysis, these technologies are only beneficial to the extent that they improve decision making in the care of patients and in the operation of a healthcare system. The reminder of the paper is structured as follows: in Section 2 a review of IoT, BDA and resource management planning as they have been applied to industry to date is presented. Section 3 provides a model of an implementation of these technologies for the most fundamental decision support in healthcare. From this model, in the Conclusions we are able to raise several issues about the challenge of making valuable use of IoT and BDA.

2. A brief literature review on IoT, BDA, hospital resource planning and value co-creation

2.1 Internet of Things

IoT is a worldwide network of interconnected objects that are uniquely addressable based on standard communication protocols. Any object (computers, mobile phones, RFID tagged devices, and especially Wireless Sensor Networks -WSN, etc.) is able to dynamically join the network, collaborate and cooperate efficiently to achieve different tasks. IoT devices gather and share information directly with each other and the cloud, enabling a fast and accurate collection, recording and analysis of new data streams (Evans, 2011).

IoT is a concept reflecting a connected set of anyone, anything, anytime, anyplace, any service, and any network. It is a megatrend in next-generation technologies that can impact the whole business spectrum since the advanced connectivity of these devices, systems, and services can go beyond machine-to-machine (M2M) scenarios (Höller et al., 2014).

In the field of healthcare, emerging solutions are WSN, dispersed and autonomous sensing stations, and the Wireless Body Sensor Network (WBSN), in which different bio-medical sensors are employed simultaneously on the patient.

Naqishbandi et al. (2015) reported that the emergence of IoT in healthcare has a dual purpose:

- Monitoring patient health over the time enabling preventive care and prompt diagnosis of complications;
- Avoiding manual collection of data (reducing the risk of errors) and making it available where needed.

In any case, there is also a non-clinical reason for IoT implementations -- managing a hospital by means of real time/ updated and detailed data. Some studies in this field are reported in literature. For example, a NFC reader was installed in each closet to track the stock level of drugs (provided by NFC tag) in a hospital (Rico et al, 2012). A more complex system was proposed to manage a medical nursing system (Huang and Cheng, 2014). This solution integrated an identity management system to identify the patient each time is needed, an environmental sensing system to control temperature, humidity and light, a biomedical system to collect and store results of health check, a medication system to control drugs administration and reordering and a personal orientation system to alert the hospital in case of patient escape. In another case, patient flow in a department was tracked and instructions on what to do and where to go were provided by means of a smartphone or smart card in order to schedule patients, better manage capacity and improve efficiency and quality of services (Chen et al., 2016).

A similar objective was pursued by D'Souza et al. (2011), who presented a Real Time Location System (and its technical characteristics) to manage disaster medical response. The solution was used to track patients and high-value equipment (like infusions pumps) automatically, collecting location data in real time. This is particularly useful in case of disaster, when there is a peak of patient demand and human resources could be scarce. For this reason, the visual management of patients is shared with other hospitals by means of a common platform, in order to eventually plan transfers. Moreover, the system recorded time spent by patients at each step of their flow in the hospital, so improvements in personnel allocations and care processes could be evaluated by governance.

Islam et al. (2015) stated that IoT can support the efficient scheduling of healthcare resources maximizing the service level to patients, but no detailed study was available. Naqishbandi et al. (2015) identified three categories of IoT devices, which we can see are relevant to healthcare service systems:

- Personal monitoring devices (as wearable blood pressure monitor, movement monitor, etc.);
• Smart home devices (monitoring living conditions at home, like temperature or lights, and keeping track of eaten food);
• Environment monitoring sensors (monitoring living conditions of cities, used in smart cities to infer on the spread of diseases like asthma).

2.2 Big Data Analytics
The term analytics is used to represent the process that extracts value from data through the creation and distribution of reports, modelling of statistics and data-mining, exploration and visualization of data, sense-making or other similar techniques (Grossman and Siegel, 2014).

BDA is the analysis of massive amounts of data of different data sources and type that can help stakeholders personalize care, engage patients, reduce variability and costs, improve quality of health delivery and also contribute to providing a rich context to shape many areas of health care like diagnosis, side-effects of drugs, genome analysis etc. (TDWI, 2011).

For the different modelling purposes of BDA, models for predictive analytics can make sense out of a large collection of data leveraging the correlations in longitudinal records into useful information for decision support or handling large volumes of medical imaging data, extracting potentially useful information and biomarkers. Other applications are currently related to the analysis of genomic data, combined with standard clinical data to provide a deeper knowledge to physicians (Naqishbandi et al., 2015).

Moreover, analytics can also help remove inefficiencies supporting hospitals in resource management by reducing emergency waiting times, track patient movements, moderate X-ray dosage levels etc. (Naqishbandi et al., 2015).

2.3 Industry 4.0

Industry 4.0 (I4.0) is a combination of the Computer Integrated Manufacturing (CIM) and Lean Management, which incorporates the benefits of both (Schuh & Stich, 2014) and is mainly driven by IoT. (Yang et al., 2014). Modern plants consist of various machines and industrial manufacturing equipment which are more or less smart. I4.0 tries to integrate their environment vertically and horizontally in order to exchange information by means of cyber-physical systems (CPS), which consists of embedded systems with the ability to communicate, preferably via internet technologies like web services (Schuh et al., 2014).

Such systems enable the adoption of real-time scheduling, which is based on a backlog-free planning algorithm for scheduling customer orders in accordance with finite capacities. In contrast to scheduling methods with assumed infinite capacities, backlog-free means that the scheduler cannot violate capacity constraints and every schedule that the planner produces is feasible. A scheduler for I4.0 CPS should be event-triggered rather than manually provoked or on a scheduled basis (e.g. every day or hour). This means that the scheduler reacts dynamically to changes and adjusts to the real situation. When a machine brakes down for example, it sends this information via web service interface. The human scheduler can see deviations for orders, planned throughput times, problems and bottlenecks, and can make changes to the plans by what-if-scenarios. In a case study, with a CPS-planning service, the waiting time was significantly reduced with the rescheduling. In consequence, the manufacturing cost saving accounts for 65,384 EUR (Berger et al., 2016). These same technologies and scheduling models can be applied, with some modifications, to healthcare systems. Figures 2 and 3 sketch this application of I4.0 to healthcare.
### 2.4 Resources and decision to engage

The concept of resource-integrating actors, promoted within the Service Dominant logic by Vargo and Lusch (2008), highlights the idea of actors that own or can access resources that allow them to operate an integration established for service-for-service exchange and value co-creation for mutual benefit. Interaction or engagement of actors to create value have been seen as the foundation of service systems (Maglio et al., 2009).

According to Lusch and Vargo (2014), actors can be humans or collections of humans, such as organizations. Moreover, researchers on sociomateriality view the human and social dimension interwoven with the realm of the material, including technologies (Orlikowski and Scott, 2008). Finally, Badinelli (2016) accommodated the dichotomy between human and non-human participants in the service system by making a distinction between agents and actors.

Then, a patient-actor can have access to resources such as family, friends, knowledge about diseases or time to be cured, while medical clinic-actor owns or can access equipment, beds, operating rooms, physicians, nurses, knowledge to cure diseases etc. Even technologies can represent agents or resources of actors depending on the lens used to analyse the service system. Indeed, advances in autonomous technologies provide the opportunity for re-shaping actor-to-actor interaction, for example, by substituting human-based...

---

**Figure 2: I4.0 use Cases in healthcare (Hosseini, 2015)**

<table>
<thead>
<tr>
<th>Digital data</th>
<th>Automation</th>
<th>Interconnectivity</th>
<th>Digital customer interface</th>
</tr>
</thead>
<tbody>
<tr>
<td>Suppliers</td>
<td>3D-printing (e.g., artificial limbs and implants)</td>
<td>Remote surgery</td>
<td>E-commerce portals</td>
</tr>
<tr>
<td>Manufacturers (Diagnostics)</td>
<td>Additive manufacturing</td>
<td>Remotely monitored implants</td>
<td></td>
</tr>
<tr>
<td>Manufacturers (Therapy)</td>
<td>Digitalization in operating theatres</td>
<td>Hybrid operations</td>
<td></td>
</tr>
<tr>
<td>Hospitals &amp; Doctors</td>
<td></td>
<td>Centralized access to health data</td>
<td></td>
</tr>
<tr>
<td>Maintenance &amp; Service</td>
<td></td>
<td>Remote maintenance</td>
<td></td>
</tr>
</tbody>
</table>

**Figure 3: Healthcare Market 2020 (Hosseini, 2015)**

- **PHYSICIAN**
  - Deciding on treatment based on real-life experience provided by advisors & increased networks
- **B2B ADVISORS**
  - Giving individual treatment advice based on real-life data and diagnostic results
  - Real-time decision support
  - Providing personal tools to diagnose conditions
  - Providing companion diagnostics as instructed by data interpreters
- **LAB**
  - Providing diagnostic services where required
- **TRADITIONAL DIAGNOSTICS**
  - Providing instruments/ tools to diagnose conditions
  - Providing companion diagnostics as instructed by data interpreters
- **DATA COLLECTORS/ DATA CONNECTORS**
  - Collecting and connecting health-related data from all sources
  - Providing data/ information to other players to increase value
- **PHARMA**
  - Developing and providing innovative drugs
  - Jointly developing companion tests with diagnostics as instructed by data interpreters
- **DATA INTERPRETERS**
  - Extracting meaningful information out of massive data sets that they may or may not own themselves (e.g., for companion diagnostic/ treatment based on real-life data)
interaction with technology-to-technology interaction, technology to human interaction or big data and advanced analytics (Brynjolfsson and McAfee, 2012).

The prerequisite for resource integration is the decision of actors to engage in the service exchange, which in the viewpoint of a patient accessing a medical clinic, is the decision to go/not go to the clinic for a visit. Since value is uniquely determined by the beneficiary (10th foundational premise of S-D logic, Vargo and Lusch, 2008), resources are not inherently valuable but they become more or less valuable depending on the context of their integration (Chandler and Vargo, 2011). The same resource - a bed in a hospital, for example - has different value for a patient needing care and a nurse manager, who knows the requirements of other patients needing a bed in the same medical unit, its current load and turnover rate.

This difference in the assessment of resources can be repeated in the assessment of the engagement decisions to integrate resources depending on the context of the focused actor. While several resource management frameworks have been proposed to manage hospital resources in order to maximize the efficiency of care or to minimize the waiting time for patients (Hans et al., 2012; Roth and Van Dierdonck, 1995), the actual benefits of these actors are usually not simultaneously maximized. This asymmetry may be further complicated by the introduction of IoT and BDA into the prediction of an ailment and then into the prediction of the commitment of the resources needed for care. In the next section we analyze these benefits, highlighting the context, interaction, and engagement required for effective resource integration and value creation.

3. A Model of Decision Support in Service

Returning to our main purpose of examining how IoT and BDA can improve healthcare service, we are compelled to model the role of these technologies in the specific case of decision support of a patient during the service journey. Only if these technologies can improve the performance of these decisions can they become valuable and worth implementing.

3.1 The Healthcare Service Journey

The healthcare service system enables the fundamental sequence of diagnosis, treatment and return to health. Apart from the patient pathway, which describes the standard sequence of activities that a patient should undertake to diagnoses and care of a certain condition or disease, we introduce the healthcare service journey. This journey, which is an expanded view of customer journey maps (Richardson, 2010), is guided by a sequence of decisions made by actors (both patients and the medical clinic in the Healthcare Service System - HSS) co-creating value. We refer to these decisions as engagement decisions because their purpose is to choose whether to engage in a service activity, which requires commitments of resources. The co-creative aspect of the HSS is the result of the joint decision to engage in a service activity and the commitment of respective resources to the service-for-service exchange. The engagement decisions are the dynamic activation mechanisms that determine the trajectory of the service journey and ultimately determine the viability of the HSS (Badinelli, 2016).

To be more specific, in an HSS the patient tracks his or her health condition through some kind of data collection and predictive analysis from which an engagement decision is made. This data collection and predictive analysis can be as basic as the patient waiting until some pain or discomfort manifests itself (the way most people currently track their health) or as intense as regular visits to a clinic for examinations by a physician. Between these two extremes, there is the possibility of an IoT-enabled tracking of health conditions at regular intervals, or even on a near real-time basis, combined with BDA to produce predictions. Regardless of the method used for data collection and predictive analysis, when the patient receives an indication of an ailment that is strong enough to motivate engagement with the HSS, the patient makes an appointment with the clinic in order to obtain a rigorous diagnosis. The signal or indication that the patient receives must exceed a certain threshold to motivate this engagement. This initial service activity is that of a careful examination and diagnosis by a physician. If the diagnosis positively identifies the presence of a condition or disease, the patient and the HSS engage in service activities for treatment.

Figure 4 shows the co-creation of value in a HSS. Both Actors (the patient and the medical clinic) provide data from IoT and other sources. The patient is the receiver of an indication of an ailment (prediction) and makes the decision to engage. This scenario is one that is enabled by IoT in the form of patient monitoring.
devices that provide a constant stream of data to an automated diagnostic model based on BDA. This information processing is the first step in the process of dynamic decision making in the service journey. An alternative scenario would send the prediction to the clinic, from which the clinic would propose an engagement with the patient by recommending an appointment for a visit.

If both actors agree to engage, they commit the respective needed available resources (family, friends, knowledge, drugs, physicians, equipment, etc.) and integrate their resources. The resource application of each actor is the service he/she exchanges with the other actor for mutual benefit and value co-creation. It, in turn, provides new and updated data related to the health status of the patient (results of exams, reactions, etc.), closing the loop.

Figure 4: Value co-creation in Health Service System (actor B, who can be the patient or the medical clinic, is the receiver of the prediction who make the decision to engage)

3.2 Decision strategies and data sources

A patient or the medical clinic initiates the decision to engage on the basis of a decision strategy, which integrates a predictive model with a decision model. Current practice manifests several types of such decisions strategies (DS). For the comparative analysis in this paper, we identify three strategies:

- **DS1**: This decision strategy is based on regular visits to a clinic for examinations (prevention). In this case, the patient engages with the medical clinic in frequent and regular examinations in order to detect the possibility of an ailment at the earliest opportunity. The predictive model of this strategy is constructed from the raw data of the vital signs and other measurements of the patient that are recorded by the clinic and the interpretation of these data by the examining physician;

- **DS2**: This decision strategy is based on IoT tracking of health conditions at regular intervals, or even on a near real-time basis. Combined with data coming from Electronic Medical Records or other sources, the predictive model requires BDA in the form of a statistical model for binary classification.

- **DS3**: This decision strategy is based on the patient waiting until a pain or discomfort manifests itself. This is the way most people currently track their health. The predictive model is constructed from the raw data of the patient’s own sensing of symptoms and the interpretation of these symptoms by the patient.
3.3 The role of prediction

It is important to notice that decision making in the service journey within the HSS is enabled by feedback control systems. The collection of data, the updating of predictive models and the review of the engagement decision are done recursively by all actors. At each stage of the healthcare service journey, the decision maker forecasts the next stage of the proposed service. With each recursion, the decision maker must estimate the outcome of any proposed engagement. This requires estimating the required resource commitments, output resource production and the value gains from the proposed service activity. Given the updated estimates, the actor can make a decision about whether or not to join the engagement. Hence, in each cycle of the feedback control system, the actor acquires data, updates estimates of the relevant parameters of the engagement and then adapts the engagement decision to the new estimated state. Hence, IoT and BDA are platforms that empower actors to co-create value within the healthcare service journey.

Prediction informs the engagement decision. Through prediction, the decision maker can estimate and forecast the outcomes of the options of either engaging the HSS or of deferring this engagement. Moreover, the quality of predictions has a direct effect on the performance of the engagement decisions. In the case under study in this paper, the performance measures that determine the engagement decision are cost of diagnosis, cost of treatment and the cost of not treating an existing condition/illness. When the patient contemplates the option of visiting the clinic for the purpose of acquiring and accurate diagnosis and, if necessary, treatment versus the option of deferring such an engagement, the key consideration is the likelihood that the patient is afflicted with an ailment that should be treated. Deferring the engagement allows the patient to avoid the cost, discomfort and inconvenience of a visit to the clinic in favor of waiting to see if later predictions increase the likelihood of the ailment to a level that warrants intervention.

From the patient point of view, when he/she contemplates the option of visiting the clinic for the purpose of acquiring and accurate diagnosis and, if necessary, treatment versus the option of deferring such an engagement, the key consideration is the likelihood to be afflicted with an ailment that should be treated. Deferring the engagement allows the patient to avoid the cost, discomfort and inconvenience of a visit to the clinic in favor of waiting to see if later predictions increase the likelihood of the ailment to a level that warrants intervention.

Given the stochastic, and often fuzzy, understanding of service engagements, these estimations can be quite inaccurate (Badinelli, 2012; Badinelli, 2016). For this reason, satisfaction with service engagements varies widely across actors and across time. Predictive analytics, enabled by the introduction of IoT, has introduced new methods for making these predictions and, it is hoped, more accurate estimates. However, these new technologies and methods bring new challenges to the design, management and control of the service systems that they support. We highlight some of these challenges with a simple model of a HSS.

3.4 A model of actor’s engagement decision

Consider a simple HSS that consists of a patient and a medical clinic. Using all available data and a predictive model, the system generates a prediction about the presence of a particular ailment. This type of prediction is produced by a model with a binary dependent variable, such as the very commonly used logistic regression model. Based on the prediction, the patient or the medical clinic (depending on who receives the first signal from the predictive model) must decide whether to make an appointment at the clinic. This decision for the patients would imply tradeoffs among cost, time, pain and discomfort, long-term health and other personal factors. In either case, the decision maker must weigh the potential outcomes of the decision by the likelihood that the prediction is correct.

The linchpin of the decision is the probability of the ailment, which depends on the availability of data and the accuracy of the predictive model. The potential errors in predicting a binary dependent variable are illustrated by a confusion matrix in Table 1.
Table 1: Confusion matrix

<table>
<thead>
<tr>
<th>Predicted Positive (reject null hypothesis)</th>
<th>Predicted Negative (fail to reject null hypothesis)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actual Positive (P)</td>
<td>TP</td>
</tr>
<tr>
<td></td>
<td>FN (Type II error)</td>
</tr>
<tr>
<td></td>
<td>P = TP + FN</td>
</tr>
<tr>
<td>Actual Negative (N)</td>
<td>FP (Type I error)</td>
</tr>
<tr>
<td></td>
<td>TN</td>
</tr>
<tr>
<td></td>
<td>N = TN + FP</td>
</tr>
<tr>
<td></td>
<td>TP + FP</td>
</tr>
<tr>
<td></td>
<td>FN + TN</td>
</tr>
<tr>
<td></td>
<td>Totals</td>
</tr>
</tbody>
</table>

Note: TP = true positive, TN = true negative, FP = false positive, FN = false negative

Binary classification models generate a probability for the presence of an ailment. A simple decision rule, in the form of a threshold ($p_c$) for this probability ($p$), is applied to classify the prediction into one of two possible outcomes – positive prediction or negative prediction.

First, we will take the point of view of the patient in making the engagement decision. We assume that the patient and the HSS update the database of relevant data and re-generate the predictive model at certain time intervals, which we call epochs. At each epoch, the engagement decision is made by the patient. As the outcomes of the decision depend on the correctness of a prediction, the rational decision maker will base the decision on a risk analysis, which follows the familiar newsvendor problem. For the patient ($P$), the parameters of this model are as follows:

$c_{PD} = \text{cost to the patient of an accurate diagnosis performed at the clinic.}$ This cost includes the monetary expense to the patient of performing tests at the clinic in order to ascertain the correct diagnosis as well as the “cost” of the inconvenience, discomfort and anxiety that is associated with the visit to the clinic.

$c_{PT} = \text{cost to the patient of treatment of the ailment after a true positive diagnosis.}$ This cost includes the monetary expense to the patient of the treatment as well as the inconvenience, discomfort and anxiety that is associated with the treatment.

$c_{PI} = \text{cost of ignoring the presence of the ailment until a more reliable diagnosis emerges.}$ This cost includes the pain, discomfort and worsened medical condition of the patient that results from delaying treatment. The magnitude of this cost depends on the nature of the condition or disease and the length of time that treatment is delayed. In fact, this cost includes the “cost” of shortened life span that could result from delayed treatment of a serious condition. In other words, the cost of ignoring the presence of a condition or disease is the increase in the cost of the treatment that will result by delaying treatment until the next epoch (which has been proposed several times in literature with different formulation in order to prioritize patient access to hospital services, Paulussen et al. 2006).

Each of these “costs” must be estimated in consideration of the actual monetary expenses of the patient and the HSS as well as the pain, discomfort, time spent, anxiety and other subjective dimensions of value. Consequently, the relative scale of these parameters can vary greatly across different conditions and diseases.

Each decision maker is faced with a rather simple choice, which we illustrate with the decision tree in Figure 5. Define, $p = \text{Precision} = \text{the probability of a true positive prediction, given a positive prediction} = \frac{TP}{P} = \frac{TP}{TP + FP}$
3.4.1 Patient’s decision rule

If the predictive model makes a positive prediction then, of the four possible outcomes identified in Table 1 and represented in Figure 5, only two are relevant to the engagement decision. Either the prediction of an ailment is correct or it is not correct. Therefore, the probability of a true positive is the key output of the predictive model. We derive the expected cost of each alternative as follows:

\[
p(c_{pl} - c_{pt} - c_{pd}) - (1-p)c_{pd} = \text{expected benefit of engaging the medical clinic}
\]

\[-pc_{pl} = \text{expected benefit of not engaging the medical clinic}
\]

The decision maker will choose to engage if,

\[
p(c_{pl} - c_{pt} - c_{pd}) - (1-p)c_{pd} > -pc_{pl}
\]

This condition implies the following decision rule for accepting the engagement proposal:

If \( p > p_c = \frac{c_{pd}}{2c_{pl}-c_{pt}} \) then engage, otherwise wait until the next epoch. \( (1a) \)

A corresponding analysis of the case of a negative prediction results in the decision rule,

If \( q > p_c = \frac{c_{pd}}{2c_{pl}-c_{pt}} \) then engage, otherwise wait until the next epoch. \( (1b) \)

where \( q = \text{False Omission Rate} = \text{probability of a false negative, given a negative prediction} = \frac{FN}{N} = \frac{FN}{(TN+FN)} \). The False Omission Rate tends to be monotonically related to the Precision.

The predictive model converts raw data about the state and history of the patient into a probability of the patient being afflicted with an ailment that warrants treatment. The probability of a true positive prediction, which we denote \( p \), can be compared with the critical probability \( p_c \) in (1a) and provides a measure of the Precision (a.k.a. Positive Predictive Value) of the predictive model that is used by the decision maker. Similarly, this critical probability determines the decision when there is a negative prediction according to (1b). Patient can use Decision Strategies 1, 2 and 3 under condition 1 to assume the engagement decision.

A comparison of the performance of these three patient’s decision strategies (DS1-3) requires an assessment of the relative precision of the predictive models on which they are based. To this assessment, we introduce a model of precision. Define,

\( a(t) = \text{the level or extent of an ailment at time } t. \) We assume that this function is increasing in \( t \), reflecting the fact that ailment worsens over time if it is untreated.
It stands to reason that the precision of any predictive model increases as the ailment progresses and becomes worse. Each method for data collection and predictive analytics has a precision for producing a true positive probability. However, for all methods, as the level of the patient’s ailment increases, this predicted probability increases. Therefore, we assume that the probability \( p \) is increasing in \( a(t) \) and, consequently, is increasing in time. Furthermore, we can reasonably assume that the predictive model with the highest precision is that of \( DS1 \) and the predictive model with the lowest precision is that of \( DS3 \), with the IoT/BDA-based predictive model performing somewhere between these extremes. Define,

\[
a_c = \text{the level of } a(t) \text{ at which the predictive model attains } p_c.
\]

\( l_1, l_2, l_3 = \) the lead time at which the predicted probability of an ailment motivates engagement, under decision strategies \( DS1, DS2 \) and \( DS3 \), respectively.

Given the relative sensitivities of the three predictive models, we can conclude that,

\[ l_1 < l_2 < l_3 \]

Figure 6 illustrates how these lead times are produced from the sensitivities of the three decision rules and the progress of an ailment.

At this point in the analysis, we should examine the cost model more closely. It is reasonable to assume that the cost to treat an ailment is increasing in the progress of the ailment. In other words, \( c_T \) is increasing in \( a(t) \), and, consequently, \( c_T \) is increasing in time. We also note that the marginal cost of deferring an engagement is the increase in the cost of treatment that would be incurred if the engagement is deferred to the next epoch. That is,

\[
c_I(t) = \frac{dc_T}{dt} \Delta t + f \left( \frac{da}{dt} \Delta t \right)
\]

where \( \Delta t = \) the length of an epoch and \( f = \) pain, discomfort, damage to organs and shortened lifespan that can be expected if the ailment is left untreated for another epoch.

We can see how different ailments present a wide range of characteristics of these cost functions. In the case of serious ailments, \( c_T \) is convex and, if the ailment is life threatening or seriously debilitating, \( f \) is convex and increases to a singularity. By contrast, in the case of mild ailments that may be little more than discomforting, the rate of increase in \( c_T \) and \( c_I \) over time may be rather shallow. In any case, we should assume that the denominator in (1) is increasing in time, which implies that the critical probability, \( p_c \), is
decreasing in time. Therefore, the lower precisions of \( DS2 \) and \( DS4 \) are somewhat offset by this decrease in the threshold for motivating an engagement. Nevertheless, condition (2) always holds.

We now turn our attention to the long-run, aggregate performance of the HSS under the three decision rules. These costs are functions of the frequency of clinic visits for examinations, the frequency of visits for diagnoses and the frequency and timing of treatments. We note that decision rule \( DS1 \) maximizes the frequency of visits to the clinic for examinations and diagnoses, but minimizes the cost of treatment. Decision strategy \( DS3 \) minimizes the frequency of clinic visits but maximizes the occurrences of delayed treatments. The performance of decision strategy \( DS2 \), the modern approach based on IoT and BDA, is better than the performance of \( DS3 \) in terms of delayed treatment but worse than \( DS1 \) in terms of diagnosis cost and treatments. \( DS3 \) performs better than \( DR2 \) in terms of diagnosis cost but worse than \( DS2 \) in terms of delayed treatment. Table 2 summarizes these comparisons.

<table>
<thead>
<tr>
<th>Decision Strategy</th>
<th>Frequency of Visits</th>
<th>Patient Expenditures on Diagnoses</th>
<th>Patient Expenditures on Treatments</th>
<th>Patient Expenditures due to Delayed Treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td>DS 1</td>
<td>High</td>
<td>High</td>
<td>Low</td>
<td>Low</td>
</tr>
<tr>
<td>DS 2</td>
<td>Medium</td>
<td>Medium</td>
<td>Medium</td>
<td>Medium</td>
</tr>
<tr>
<td>DS 3</td>
<td>Low</td>
<td>Low</td>
<td>High</td>
<td>High</td>
</tr>
</tbody>
</table>

### 3.4.2 Medical Clinic’s decision strategy

As we explained earlier, the engagement can be proposed by either the patient or the clinic. If the clinic pondered the engagement decision, the tradeoffs involve the same performance measures of the decision, but the values of the parameters are different.

The HSS includes the supply chain for all medications, equipment and medical personnel that are required for diagnosis and treatment. The acquisition of these resources requires lead times that are much longer than the time that a patient can wait for a treatment (Iannone et al., 2011). The medical clinic, like all service systems must have capacities and inventories of these resources in place to provide service as it is demanded, with an acceptably low wait time for the patient. Consequently, the cost parameters of the medical clinic have a specific nature (Iannone et al., 2013) and can be represented by the allocation of the overhead costs of maintaining the HSS to specific procedures for diagnosis or treatment. If we assume that the Medical Clinic is committed to maintaining a certain level of service, then we must model these costs as increasing functions of the randomness in patient demand due to the phenomenon of service systems that we explain below.

It is a well-established fact that the distribution of wait times of a service system is a function of the capacity of the service system, the average rate of demand for service and the volatility of the demand rate and of the service process times. Figure 7 shows the familiar tradeoff between wait time and system capacity for a simple, single-queue service system with general distributions for inter-arrival times and process times (Kingman 1961). Numerous theoretical and empirical studies have borne out this kind of relationship over every conceivable kind of service system. The phenomenon that is illustrated in Figure 7 is the effect of randomness on the performance of a service system. The wait time function is convex. Furthermore, the rate by which the wait time increases as a function of traffic intensity increases with the amount of randomness in the arrival times of patients and in the process times for medical procedures. Service systems that experience higher unpredictability in demand rates and/or in service process times experience higher wait times. To model this phenomenon in what follows, we will express the cost parameters of the HSS as functions of the uncertainty of the demand process. Given that \( M \) stands for medical clinic, we define,

\[
\lambda = \text{the average rate of arrivals of patients to the clinic}
\]

\[
\sigma = \text{standard deviation of the time between arrivals of patients to the clinic}
\]

The corresponding cost parameters for the medical clinic are as follows:
\[ c_{MD}(\lambda, \sigma) = \text{cost to the medical clinic of diagnosis in the form of the expense of performing tests at the clinic to ascertain the correct diagnosis} \]

\[ c_{MT}(\sigma) = \text{cost to the medical clinic of treatment of the ailment} \]

\[ c_{MI}(\sigma) = \text{cost to the medical clinic of ignoring the prediction and delaying treatment of the ailment} \]

Now, let us consider the expenditures of the medical clinic. It must provide the resources to accommodate the demand for diagnoses and treatments that are generated by the DSs that govern the engagements. Each engagement DS affects the performance of the medical clinic due to the pattern of demand for healthcare services that the rule engenders. Therefore, in order to maintain the capacities and inventories that are required to maintain a given level of service, the three cost parameters of the medical clinic are increasing functions of the frequency of visits, \( \lambda \), and the volatility of the time between visits, \( \sigma \).

![Figure 7: The service level to capacity tradeoff](image)

\( c_{MT}(\sigma), c_{MI}(\sigma) \) are increasing functions of the randomness in patient arrivals

\( c_{MD}(\lambda, \sigma) \) is an increasing function of frequency of visits and randomness in patient arrivals.

The three decision strategies applied by the patient generate different demand patterns in terms of the randomness of the demand for healthcare services and the extent of the services demanded. Using the results from Table 2, we derive the comparison of performance of the medical clinic under the three decision rules, which is shown in Table 3.

<table>
<thead>
<tr>
<th>Decision Strategy</th>
<th>Frequency of Visits</th>
<th>Uncertainty in Timing of Visits</th>
<th>Medical Clinic Overhead Expenditures for Diagnoses</th>
<th>Medical Clinic Overhead Expenditures for Treatments</th>
<th>Medical Clinic Overhead Expenditures due to Delayed Treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td>DS 1</td>
<td>High</td>
<td>Low</td>
<td>High</td>
<td>Low</td>
<td>Low</td>
</tr>
<tr>
<td>DS 2</td>
<td>Medium</td>
<td>High</td>
<td>Medium</td>
<td>Medium</td>
<td>Medium</td>
</tr>
<tr>
<td>DS 3</td>
<td>Low</td>
<td>High</td>
<td>Low</td>
<td>High</td>
<td>High</td>
</tr>
</tbody>
</table>

Table 3: Long-run expenditures of medical clinic under different decision rules
A final note about the medical clinic decision is how the difference in the parameters of the decision for the medical clinic lead to a different optimal intervention plan from the one that the patient would choose. A similar analysis of the engagement decision of the medical clinic produces the following decision rule:

If \( p > p_c = \frac{c_{gg}(\lambda, \sigma)}{2c_{gg}(\sigma) - c_{gg}(\sigma)} \) then engage, otherwise wait until the next epoch.

The asymmetry in the cost parameters between the patient’s decision and the medical clinic decision imply that the patient and the medical clinic have different thresholds for deciding when a patient should visit the clinic. This conflict will manifest itself in one participant or the other failing to optimize its costs.

4. Conclusions

Our analysis of the decision rules for the engagement decisions reveals some interesting insights. First, we can offer a critique of the effectiveness of IoT and BDA for predictive modeling in support of healthcare decisions.

1. On the asymmetric effects of decisions: It is generally believed that the use of IoT and BDA for predictive modeling will enable a form of demand management that will benefit all participants in the service system. However, the use of these technologies is not a win-win proposition for the patient and the medical clinic. The use of IoT and BDA for predictive analytics in support of the engagement decisions has asymmetric effects. Benefits that accrue to the patient do not transfer to the medical clinic, which, in turn, may find its costs increase as a result of the patient adopting decision strategy \( DS_2 \) over the current most common decision rule \( DS_3 \). Furthermore, this phenomenon is general, in the sense that it applies to most service systems. The root cause of the inability of the medical clinic to capture benefits from the use of IoT and BDA stems from a fundamental characteristic of service systems – the inertia of the service system constrains the rate of change of capacity and inventory. These constraints exist in all service systems. Changes to the decision strategy (1) will be necessary to implement Demand Response for load-leveling of the medical clinic.

2. On the effectiveness of BDA predictions: The value of IoT in healthcare can be realized only to the extent that it is integrated with BDA that is sophisticated enough to improve substantially the probability estimates that determine the engagement decision according to (1). The precision of the predictive models that are available to the patient has a profound effect on the performance on the health and expenses of the patient and on the cost of maintaining the medical clinic.

3. On the flexibility of resource capacity and scheduling of medical clinic: The ability of a medical clinic to respond to the anticipated improvement in the detection of ailments from the use of IoT and BDA is limited by the natural inertia of capacitated service systems to respond to demand. Therefore, as the adoption of these technologies raises expectations of prompt intervention, there will be increasing pressure on medical clinic to implement lean practices, process improvements and finite capacity scheduling that will enhance the ability of the medical clinic to respond to demand volatility. Given the limitations and expense of improvements, we can assert that ultimately, patient and medical clinic will have to cooperate to adjust the scheduling of diagnoses and treatments in consideration of both patient costs and medical clinic costs, which will require combinations of advanced scheduling or delayed scheduling of some interventions in order to compromise between patient desires and medical clinic overhead costs. In this way, the healthcare system becomes truly co-creative. Fortunately, IoT and BDA provide both sets of actors with the information that they need to create this service innovation.

There are many new challenges and issues being raised by the introduction of IoT and BDA as tools for smart healthcare service systems. This paper has only introduced some of the most basic questions about this burgeoning trend in healthcare. Much research remains to be done. Of immediate interest is the estimation of the cost parameters that are defined in this paper for different kinds of ailments in order to study the effects on decision making of the variability of these estimates across patients. Another very interesting challenge that the use of IoT and BDA presents is that of differential diagnosis. As patients receive more information
about potential ailments from sophisticated and data-driven predictive models, there will emerge a challenge in the form of distinguishing symptoms of serious ailments from those of minor ailments. Already, physicians have noticed the syndrome of “cyberchondria” among some patients. Consequently, more data will not necessarily lead to a more efficient and effective HSS.
References


