Clinical Decision Support for Integrated Cyber-Physical Systems: A Mixed Methods Approach

Alex Roederer
Department of Computer Science
University of Pennsylvania
Philadelphia, PA
roederer@seas.upenn.edu

Andrew Hicks
Department of Bioengineering
University of Pennsylvania
Philadelphia, PA
Andrew.Hicks@uphs.upenn.edu

Enny Oyeniran
Hospital of the University of Pennsylvania
Philadelphia, PA 19104
Enny.Oyeniran@uphs.upenn.edu

Dr. Insup Lee
Department of Computer Science
University of Pennsylvania
Philadelphia, PA
lee@seas.upenn.edu

Dr. Soojin Park
Department of Neurology
Hospital of the University of Pennsylvania
Philadelphia, PA
Soojin.Park@uphs.upenn.edu

ABSTRACT

We describe the design and implementation of a clinical decision support system for assessing risk of cerebral vasospasm in patients who have been treated for aneurysmal subarachnoid hemorrhage. We illustrate the need for such clinical decision support systems in the intensive care environment, and propose a *three pronged approach* to constructing them, which we believe presents a balanced approach to patient modeling. We illustrate the data collection process, choice and development of models, system architecture, and methodology for user interface design. We close with a description of future work, a proposed evaluation mechanism, and a description of the demo to be presented.

Categories and Subject Descriptors

H.5.2 [**Information Interfaces and Presentation (e.g., HCI)**]: User Interfaces; I.2.1 [**Artificial Intelligence**]: Applications and Expert Systems

General Terms

Design, Experimentation

Keywords

Clinical Decision Support, Vasospasm, Machine Learning, Mixed Methods, Patient Modeling

1. INTRODUCTION

Many modern critical care units continuously monitor patient vital signs. Due to the volume of data produced and the lack of sophisticated on-line analytic tools, however, this data is currently of limited use. Bedside nurses can utilize it through threshold-based alarms, but accessing past data is often difficult. As physicians are not geographically bound to one or two bedsides, they must rely on periodic manual recording of the data on paper records, or hand-entered values in electronic medical records. There may be information within this data that can assist clinicians in real-time decision making, but it is inaccessible due to technological and human factor barriers.

To more effectively use this data, systems are being developed that enable real-time medical information from multiple devices to be gathered in one place [11]. It would then be possible to analyze this data alongside information from a patient's electronic health record to create a clinical context. Once this context is created, "smarter" alarms can be created that detect physiologic changes in the patient and present relevant pieces of information to the physician. The clinician can use the system to investigate further, probing or visualizing the data in real-time.

We propose a design of an alarm system which utilizes multiple patient models using data from different sources, forming a more complete picture of the patient's state and balancing the traditional drawbacks of complex statistical models. We describe in particular a project in which multiple patient models are used to communicate a patient's risk for vasospasm after aneurysmal subarachnoid hemorrhage.

2. BACKGROUND

Some vital sign monitors can be configured with threshold alarms, which activate when the value monitored crosses a predefined threshold. These devices can be vital for the timely detection of emergency states [1, 10], but suffer from severe limitations. They are simplistic, unable to recognize anything beyond basic threshold crossings; they are insular, only utilizing single streams of data; and they do not leverage patient context information. As a result, they produce many false positive alarms which fatigue caretakers [3, 9]. Many efforts have been made to improve the accuracy of threshold alarms [4, 12, 15] and many clinical decision sup-

port systems have been shown to hold promise in improving care [6, 8, 16]. However, many developed systems operate on hospital-dependent platforms, making use in other hospitals difficult or impossible. Additionally, adoption rate of developed systems after the main evaluation study has been completed has been low, often due to trust issues or lack of clear perceived benefit.

Changes in physiologic state can have subtle yet interconnected manifestations in phenotype, monitor, and laboratory data. Complex physiological state changes are not adequately detected by threshold-based alarms, because such detection requires integration of disparate data sources [14, 2].

Once integrated, these data sources must be used to build patient models with high accuracy, while maintaining transparency and simplicity. These models can be generated using automated statistical methods, which are able to detect patterns within data that unaided human methods may miss. However, new clinical alarms may be accompanied by a healthy amount of skepticism in clinicians, due to their complexity, their lack of transparency, and their imperfect accuracy.

3. THREE-PRONGED APPROACH

In designing a smart alarm, it is important to enable clinicians to trust statistical methods through transparency. In our *three pronged approach*, we incorporate a thorough analysis of existing clinical care guidelines, a survey of the approach commonly taken by physicians, and complex statistical models trained on data from large patient populations. Each of these approaches to clinical decision support has strengths and weaknesses, and we believe presenting them to the clinician in parallel grants each significantly more strength, and the performance of each can be evaluated against the others in order to gain a better understanding of their significance.

Clinical guidelines establish a standard of care for a particular institution and are based on an interpretation of evidence-based medicine and local expertise. Incorporating these guidelines into a decision support system which evaluates data in real-time provides a "lower bound" on behavior, and a baseline for comparing the performance of more complex classifiers. Most guidelines are relatively simple and are easily understood by humans, and incorporating them into a decision support system provides transparency and assurance of reasonable behavior.

Guidelines, however, are inherently incomplete [5]. It is impossible for every possible patient scenario to be outlined and presented, and often specificity must be sacrificed to keep the guidelines brief and easy to understand. Because guidelines are written in natural language for human readability, they may contain ambiguities, and are difficult to check for completeness. Lastly, because they are usually created by large panels of experts, they are infrequently revised and may become out of date quickly, and cannot adapt very swiftly to changes in practice suggested by new studies.

Physician experience expands considerably upon clinical practice guidelines. Physicians leverage extensive education and experience to develop and refine their own decision making techniques. Experienced physicians are able to quickly identify salient points to focus on when attempting to assess patient risk. By capturing experienced physicians' mental models in a clinical decision support system, doctors and nurses would have access to the aggregate opinion of expert consultants' opinions when no expert is available locally or immediately. More generally, they would allow patient risk to be assessed based on what an expert physician might do if they were physically at the bedside and had ample time, memory, and the ability to process large amounts of data quickly.

Utilizing physician experience is not without its difficulties. Clinicians are inherently biased, relying on past experience and practice, which is influenced by time, training, and many other factors. Also, it is often difficult to capture how clinicians actually make decisions, as they may rely on difficult to quantify "gut feelings," and may, while describing their thought process, unintentionally leave out subtle criteria that they use to accurately assess patient risk. There is also no guarantee that any group of physicians will agree: constructed models may contain contradictory statements and methods of practice in areas of clinical equipoise, which might be challenging to reconcile in an electronic system.

Electronic decision support systems are unique in that they can use statistical models learned on large quantities of relevant patient data to generate decision support. Such models can identify which features are most indicative of patient risk according to the provided data. They have the potential to identify medically novel approaches to patient risk assessment by uncovering subtle patterns in patient data normally overlooked by clinicians.

The large amount of data required to build statistical models poses significant obstacle to their use. Data collection is currently an arduous process at many hospitals, and insufficient data could lead to a poor model. Also, many statistical models act as "black boxes," producing an output but no accompanying justification, making them unintuitive and difficult to trust in clinical use. All of these present unique challenges that must be addressed for statistical data modeling to be a useful part of future systems.

4. PROJECT DEVELOPMENT DETAILS

After aneurysmal subarachnoid hemorrhage, patients are kept in the ICU for up to fourteen days to monitor for cerebral vasospasm (VSP), a narrowing of the blood vessels in the brain. VSP can lead to cerebral ischemia and neurologic dysfunction if untreated. While there are clinical factors which increase suspicion for VSP, the ability to define its onset in these patients is made difficult by poor sensitivity of available tests. The only definitive measure of VSP's presence is cerebral angiogram, which is invasive and resource-intensive. Early detection and treatment of VSP is the mainstay of ICU goals toward improving patient outcome after subarachnoid hemorrhage. With these considerations in mind, we have produced a decision support system which could aid clinicians in assessing a subarachnoid hemorrhage patient's risk for VSP during their post-surgery stay in the neurological ICU by integrating and analyzing multiple patient vital signs. We aimed to achieve a "three pronged" approach, described above, in an attempt to overcome the individual weaknesses of the various types of models.

Initial focus was to identify patient vital signs which were likely to be of use in assessing a patient's risk for VSP and could be used in the decision support system. These "features" were gathered from reviews of clinical guidelines, interviews with clinicians, and literature reviews. We also noted the current availability of these features in the neuro-

Figure 1: A portion of the statistical model testing interface, running a subset of available models over a set of patient data. The arrow points to the line indicating the time when doctors confirmed vasospasm using cerebral angiogram. Note the general upward trend of the various classifiers prior to the positive angiogram.

logic ICU at the Hospital at the University of Pennsylvania. Even if a feature was not well established as a clinical marker for vasospasm, we attempted to include it in our collection process. If we determine that these variables are not significant, their impact on the models' performance will be minimal.

As device interoperability and data accessibility are still often significant challenges in the ICU, we focused on retrospective data acquisition measures. We collected data from 89 HUP patients who presented with aneurysmal subarachnoid hemorrhages between 2001 and 2011. We gathered static data (i.e., those factors that did not vary during the patient's hospital stay) such as gender, age, tobacco and cocaine history, home medications, etc. Periodic data (taken regularly during the hospital stay) was also collected in order to establish a timeline of diagnosis and treatment of vasospasm for each patient. We tracked the results of daily transcranial Doppler tests and noted increased suspicion of vasospasm in conjunction with increased velocity of blood flow. Angiogram and CT angiogram data were used to establish definite evidence of vasospasm. We also gathered CT and MRI information, noting whether either scan evidenced stroke or neurological deficit.

To achieve the rule-based portion of the "three-pronged approach," we utilized the Hospital of the University of Pennsylvania's clinical guidelines for care of subarachnoid hemorrhage patients after surgery. We identified a section describing mechanisms for detection and treatment of VSP. Adjectives in this section were stratified, and guidelines were formulated into rule-action pairs, which were encoded as a decision table. During the encoding process, we identified several ambiguities in the rules and sought to reconcile them by consulting physicians.

The physician-based model proved to be more difficult to construct. There are many different knowledge acquisition techniques available for extracting expertise from experts, and each has strengths and weaknesses [13]. For each physician, we attempted to determine which variables they perceived as clinically relevant from among those included in the main list. We then developed a decision tree using those variables through enumeration of many possible cases. A decision tree was chosen both for its relative simplicity, ease of construction, and because it seems to correspond closely to how physicians vocalize their thought process. Then, to combine the classifiers produced, we use an "ensemble of classifiers" approach in which each physician model is run on incoming data and a mixture of their "opinions" is used as the final output, through voting or averaging.

For preliminary testing of statistical models, we utilized the Weka [7] machine learning tool. Common machine learning techniques were used to produce statistical models over collected patient data. Because it is unlikely than any single model will produce the "best" results in all scenarios, we created many different models and a testing interface 1 which can evaluate models in parallel over test data and produce detailed statistics on their performance. Results of these tests can then be used to select and combine the strongest models.

The final stage of the project involved development of a concise, intuitive user interface to present results to the user. We aimed to achieve a balance between ease of understanding, and depth of knowledge presented, by allowing clinicians to obtain more details and explore the data on their own, beyond the data summary presented to them. For the demo, we will present a simple user interface that has basic functionality aspects that illustrates this balance.

Future work will involve clinical evaluation, as well as expanding the framework developed for this system to modules for other difficult-to-diagnose conditions, such as sepsis.

5. EVALUATION

For patients who underwent an angiogram which confirmed VSP, it is possible to evaluate the system's performance in a "time-to-diagnose" capacity. In those cases where physicians performed an angiogram on the patient, we will will evaluate whether the system produced a marked increase in risk level before VSP was angiographically confirmed. We will focus on using the data available to prove that the model has a high positive predictive value, and minimal false negatives. Once this has been established, we will introduce the system into prospective trials and attempt to evaluate the system's ability to alert the physician to points of interest in the data.

6. DEMO DESCRIPTION

The overall architecture of the system will then consist of the developed statistical models to be used to evaluate the patients' state, chosen prerecorded clinical data used to train those models, input devices (in the case of the demo, prerecorded test data will be inserted into the models over time, for the purposes of simulation), the framework used to run these models in parallel and deliver their results to some display, and the display code itself, which will present the results in a manner conducive to the hospital environment.

The demo to be presented will involve an interactive vasospasm monitoring system running over patient data "received" in simulated real-time. The system will display the status of the patient and an evaluation of their risk for vasospasm. The demo audience will be able to pause the demo, move through the data forward and backward in time, in order to compare the patient's current state to their past state, and will be able to access more details about the data.

7. REFERENCES

[1] Clinical Alarms Task Force. Impact of clinical alarms on patient safety. Technical report, American College of Clinical Engineering Healthcare Technology Foundation, 2006.

[2] Clinical Alarms Task Force. Impact of clinical alarms on patient safety. *Journal of Clinical Engineering*, 32(1):22–33, 2007.

[3] J. Edworthy and E. Hellier. Alarms and human behaviour: implications for medical alarms. *British Journal of Anaesthesia*, 97:12–17, 2006.

[4] J. M. Feldman and M. H. Ebrahim. Robust sensor fusion improves heart rate estimation: Clinical evaluation. *Journal of Clinical Monitoring and Computing*, 13:379–384, 1997.

[5] M. J. Field, K. N. Lohr, and I. of Medicine U.S.. Committee on Clinical Practice Guidelines. *Guidelines for clinical practice : from development to use*. National Academy Press, 1992.

[6] A. X. Garg, N. K. J. Adhikari, H. McDonald, M. P. Rosas-Arellano, P. J. Devereaux, J. Beyene, J. Sam, and R. B. Haynes. Effects of computerized clinical decision support systems on practitioner performance and patient outcomes: A systematic review. *Journal of the American Medical Association*, 293:1223–1238, 2005.

[7] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. H. Witten. The weka data mining software: An update. *SIGKDD Explorations*, 11(1), 2009.

[8] D. L. Hunt, R. B. Haynes, S. E. Hanna, and K. Smith. Effects of computerized clinical decision support systems on physician performance and patient outcomes: A systematic review. *Journal of the American Medical Association*, 280:1339–1346, 1998.

[9] M. Imhoff and R. Fried. The crying wolf: Still crying? *Anesthesia and Analgesia*, 108(5):1382–1383, 2009.

[10] M. Imhoff and S. Kuhls. Alarm algorithms in critical care monitoring. *Anesthesia and Analgesia*, 102(5):1525–1536, 2006.

[11] A. King, S. Procter, D. Andresen, J. Hatcliff, S. Warren, W. Spees, R. P. Jetley, P. L. Jones, and S. Weininger. An open test bed for medical device integration and coordination. In *ICSE Companion*, pages 141–151. IEEE, 2009.

[12] C. Oberli, C. Saez, A. Cipriano, G. Lema, and C. Sacco. An expert system for monitor alarm integration. *Journal of Clinical Monitoring and Computing*, 15:29–35, 1999.

[13] J. R. Olson and H. H. Rueter. Extracting expertise from experts: Methods for knowledge acquisition. *Expert Systems*, 4:152–168, 1987.

[14] A. Otero, P. Felix, S. Barro, and F. Palacios. Addressing the flaws of current critical alarms: a fuzzy constraint satisfaction approach. *Artificial Intelligence in Medicine*, 47(3):219 – 238, 2009.

[15] R. Schoenberg, D. Z. Sands, and C. Safran. Making icu alarms meaningful: a comparison of traditional vs. trend-based algorithms. *Proceedings of the American Medical Informatics Association Symposium*, pages 379–383, 1999.

[16] I. Sim, P. Gorman, R. A. Greenes, R. B. Haynes, B. Kaplan, H. Lehmann, and P. C. Tang. Clinical decision support systems for the practice of evidence-based medicine. *Journal of the American Medical Informatics Association : JAMIA*, 8(6):527–534, 2001.