Incorrect or Defective Pill Detection Using a Dynamic Data-Driven Application System Paradigm

Craig C. Douglas, *Member, IAENG*, *Gundolf Haase,* Thad Hannel, David Link, and Robert A. Lodder

***Abstract*— Identification of defective tablets at the time of creation is currently infeasible. Medications are dispensed to patients in health care locations without a final check to see if a patient is receiving the correct medications. A prototype of a practical embedded, network based sensor system to solve both problems is described in this paper.**

***Index Terms*—DDDAS, integrated sensing and processing, modeling, sensor networks.**

I. Introduction

Administration of incorrect medications by professional caregivers is estimated in 1997 to have killed as many as 44,000 to 98,000 Americans after prescriptions were filled [1]. These numbers are likely to be underestimates due to unreported deaths. To put this number in perspective, use of incorrect medication is the eighth leading cause of death in the United States and actually kills more people in a given year than traffic accidents, breast cancer, or AIDS. The situation is no better in 2007.

A secondary issue is defective tablets coming off a pharmaceutical production line or mistaken packaging. Many errors are readily visible and are caught immediately. However, not all are detected and the defective or mislabeled tablets reach the marketplace.

In Section 2, we discuss the advantages of using a real-time dynamic approach instead of using static data.

In Section 3, we discuss why catching errors at the pharmaceutical production and packaging areas is essential to reducing recalls and should be part of process analytical technologies.

In Section 4, we describe an integrated acoustic sensing and processing device. A handheld version can also be used to identify medications before a caregiver delivers them to individuals. We also describe a cyber physical system (CPS) to detect incorrect or defective tablets.

Manuscript received December 30, 2008. This work was supported in part by the U.S. Department of Commerce under Grant BS123456 (sponsor and financial support acknowledgment goes here).

Craig C. Douglas is with the School of Energy Resources and the Mathematics Department, University of Wyoming, Laramie, WY 82071 USA and the Computer Science Department, Yale University, New Haven, CT 06520 USA. (phone: 307-766-6580; fax: 203-547-6273; e-mail: craig.c.douglas@gmail.com).

Gundolf Haase is with the Institute for Mathematics and Computational Sciences, Karl-Franzens University of Graz, A-8010 Graz, Austria. (e-mail: gundolf.haase@uni-graz.at).

Thad Hannel, David Link, and Robert A. Lodder are with the Chemistry Department, University of Kentucky, Lexington, KY 40506 USA. (e-mails: tshannel@gmail.com, dlink4884@gmail.com, and lodder@uky.edu).

In Section 5, we provide simple results based on a prototype system that has been built and tested in a limited manner.

In Section 6, we provide conclusions and briefly describe what needs to done next.

II. Dynamic versus static data

A data driven system allows for the implementation of real-time data to model or predict a measurement or event. By incorporating data dynamically rather than statically, the predictions and measurements become more reliable.

Consider weather forecasting. If predictions are made based on static data collected from sparsely distributed sensors, then rapidly changing conditions often make a prediction obsolete shortly after it is made. A more reliable forecasting system continuously incorporates real-time changes from many sensors into its predictions so that the forecast is always built around current conditions. As the conditions change, so does the forecast, in real-time.

Data driven applications have the ability to guide their measurement processes and refocus their resources, much as forecasts guide US Air Force 53rd Weather Reconnaissance Squadron aircraft away from calm seas and into the eyes of hurricanes to concentrate their data collection. The information collected makes possible advance warning of hurricanes and increases the accuracy of hurricane predictions and warnings by as much as 30 percent [2].

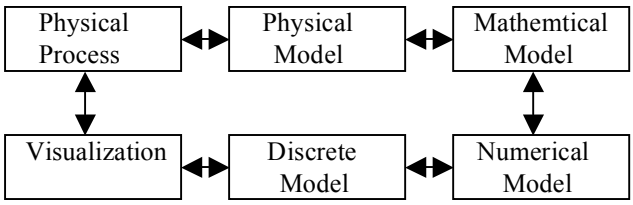
Dynamic data-driven application systems (DDDAS) provide a paradigm that is ideal for designing a network of intelligent sensors that form a symbiotic relationship with the computational model we use in this project. We quote from the 2005 DDDAS NSF solicitation, “DDDAS is a paradigm whereby application (or simulations) and measurements become a symbiotic feedback control system. DDDAS entails the ability to dynamically incorporate additional data into an executing application, and in reverse, the ability of an application to dynamically steer the measurement process. Such capabilities promise more accurate analysis and prediction, more precise controls, and more reliable outcomes. The ability of an application to control and guide the measurement process and determine when, where, and how it is best to gather additional data has itself the potential of enabling more effective measurement methodologies. Furthermore, the incorporation of dynamic inputs into an executing application invokes new system modalities and helps create application software systems that can more accurately describe real world, complex

systems. This enables the development of applications that intelligently adapt to evolving conditions and that infer new knowledge in ways that are not predetermined by the initialization parameters and initial static data. The need for such dynamic applications is already emerging in business, engineering and scientific processes, analysis, and design. Manufacturing process controls, resource management, weather and climate prediction, traffic management, systems engineering, civil engineering, geological exploration, social and behavioral modeling, cognitive measurement, and bio-sensing are examples of areas likely to benefit from DDDAS."

DDDAS environments require new software capabilities for application modeling and composition, dynamic runtime, resource management, data management, and measurement control aspects, as well software architecture drilling across all layers and end-to-end software infrastructure. The DDDAS program solicitation includes a comprehensive list of challenges and has inspired the scientific community, as exemplified by DDDAS projects that have started to address these and other related challenges. In our own DDDAS projects, we have identified several relatively diverse areas that have common issues that must be addressed by DDDAS: computer science, informational, and computational sciences, that lead to significant impact for addressing important problems. These include:

1. Effectively assimilating continuous streams of data into running simulations. These data streams most often will be…
   a. Noisy but with known statistics, and must be incorporated into the model using stochastic methods, such as filters and smoothers.
   b. Received from a large number of scattered remote locations and must therefore be assimilated to a usable computational grid.
   c. Missing bits or transmission packets, as for example is the case in wireless transmissions.
   d. Injecting dynamic and unexpected data input into the model.
   e. Limited to providing information only at specific scales, specific to each sensor type.
2. Warm restarting simulations by incorporation of the new data into parallel or distributed computations, which require the data but are sensitive to communication speeds and data quality.
3. Tracking and steering (control of measurements, models, reporting results, and visualization) of remote distributed simulations to efficiently interact with the computations and to collaborate with other researchers.
4. Translation components to rectify when simulation output parameters do not directly match observational data.
5. Interpretation and analysis components to assist researchers with collections of simulations.
6. Application program interface and middleware components for designing and creating a DDDAS or DDDAS problem solving environment.
7. Better scheduling of computational and network resources so that multiple models, possibly running at different locations, can be coordinated and data can be exchanged in a timely manner.

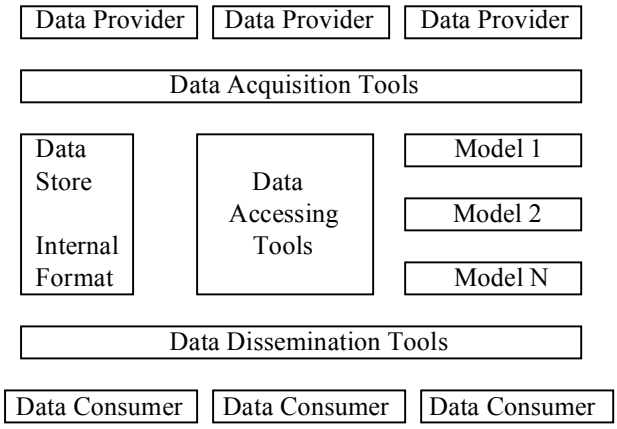
DDDAS assumes that application components, resource requirements, application mapping, interfaces and control of the measurement system can be modified during the course of the application simulation. The diagram in below shows how a number of elements might dynamically interact with each other.



Any of the components may change without resorting to a new simulation as the computation progresses. Many DDDAS applications are multiscale in nature. As the scale changes, models change, which in turn, changes which numerical algorithms must be used and possibly the discretization methods. DDDAS applications involve a complicated time dependent, nonlinear set of coupled partial differential equations, stochastic or agent-based simulation methods, which add to the complexity of dynamically changing models and numeric algorithms. It also causes computational requirements to change, particularly if dynamic adaptive grid refinement or coarsening methods are used, in response to the dynamically streamed data into the executing model.

To support data management needs in our DDDAS projects, data acquisition, data accessing, and data dissemination tools are typically used. Data acquisition tools are responsible for retrieving of the real-time or near real-time data, processing, and storing them into a common internal data store. Data accessing tools provide common data manipulation support, e.g., querying, storing, and searching, to upper level models. Data dissemination tools read data from the data store, format them based on requests from data consumers and deliver the formatted data to the data consumers.

We illustrate a simplified view of the software framework of a typical DDDAS below.



In our implementation, the data used to drive a DDDAS system are retrieved periodically by a data retrieval service, extracted, converted, quality controlled, and then staged as dynamic inputs to our simulation models. The extraction

process reads the retrieved data based on the meta data associated with them and feeds the extracted values to the conversion model whose major purpose is unit conversion, e.g., from inches to millimeters. The converted data are then analyzed for potential errors and missing values by the quality control model. This control process will ensure the correctness of the data, which is of great importance for the model simulation accuracy. The quality controlled data are then fed to the data storage model, which either saves the data to a central file system or loads them to a central database (this depends on project requirements). The data store model may also need to register the data in a metadata database so that other models can query it later.

A community web site, http://www.dddas.org, has been developed by Prof. Douglas with help from about 50 other DDDAS-related projects. The site currently has a complete funded project list (from 2000 to 2008), virtual proceedings from workshops from 2000 through 2008, a number of talks on topics that range from disaster management to transportation to homeland security to how a bat flies, news items, pointers to working DDDAS codes, and the January 2006 NSF DDDAS workshop report. Most of the projects listed are from the United States, though a number of the projects have international partners and interest in DDDAS overseas has been increasing.

III. Catching Mistakes at the Source

Numerous large pharmaceutical manufacturers outsource their small-scale manufacturing needs as a way of reducing cost or meeting their production deadlines. A contract manufacturer may make several kinds of pills that are similar in appearance at almost the same time, e.g., testing various dosages and placebos for clinical trials. A contract manufacturer may also produce pills for multiple companies. One way to reduce the possibility that pills may inadvertently become confused or contaminated is to employ a rapid and nondestructive means of verifying tablet identity. Such systems for identifying contaminated or mislabeled products must be strategically placed to prevent problems with pills before they are shipped. Process analytical technologies (PAT) on the production line should have the ability to work in real-time. Currently there are no foolproof methods to eliminate mislabeling or contamination. As a result, millions of pills are recalled in some years.

For example, in November 2006, 11 million bottles of contaminated acetaminophen were voluntarily recalled by the Perrigo Company of Allegan, Michigan due to contamination of the tablets with metal wire 3 The FDA admits that current good manufacturing processes (cGMP) have reached their limits and better "science-based" approaches are needed to insure product safety [3]. PATs are designed to prevent large recalls by detecting problems before they occur.

IV. An Integrated Sensing and Processing Approach

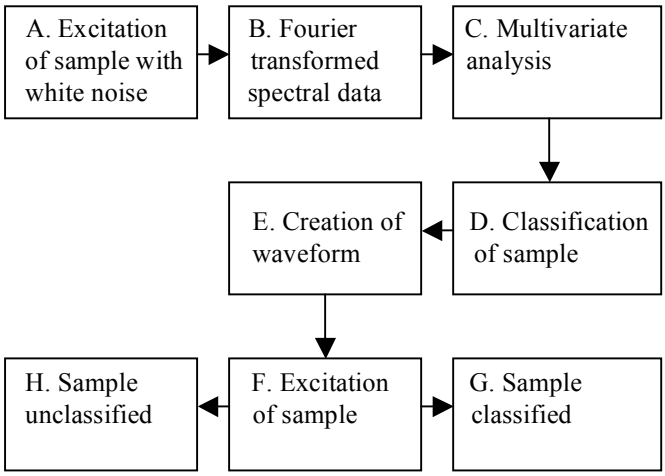
Integrated sensing and processing acoustic resonance spectroscopy (ISP-ARS) is a novel approach to conventional acoustic spectroscopic techniques. In ISP-ARS, an ISP acoustic waveform is created such that it comprises only the distinguishing spectral details associated with an analyte in question. Fourier transform acoustic resonance spectroscopy (FTARS) is used to develop ISP acoustic waveforms employed in differentiating different drugs.

ISP-ARS is fast and non-destructive. Acoustic methods are able to deeply penetrate many types of opaque packaging, in contrast to near-infrared and other optical methods. The ability to penetrate many types of packaging can be a distinct advantage in preparation of clinical trial lots, where drugs and placebos must be blinded from users. As a PAT, a series of ISP-ARS sensors could potentially scan every pill produced by a manufacturer, enabling the removal of only those pills that did not meet quality standards. A dynamic system should control a manufacturer's product line based on measurements from a series of ISP-ARS sensors, adjusting process conditions and ingredients in real time based on actual process measurements [4], [5].

ISP-ARS reduces the time required for processing that is normally observed with full spectrum FTARS. An ISP acoustic waveform is the result of chemometric analysis of the FTARS spectrum. By weighting the frequency changes according to their individual component scores, an acoustic waveform can be made that excites only those frequencies important to the analyte under observation.

The ISP output is a voltage that can be read immediately and corresponds to only the analyte under investigation.

Creation of the ISP acoustic wave begins with the chemometric analysis of the initial FTARS data. Therefore, FTARS itself makes a prediction about what will work as an ISP acoustic waveform for a given set of samples. This training process can be viewed as a dynamic data-driven application system when the performance of the ISP waveform is continuously monitored and the ISP waveform is continuously adjusted through retraining.



The traditional FTARS cycle is A→B→C→D→A. In traditional FTARS, samples are scanned and classified according to their inter-cluster distances found via multivariate analysis (A-D). This process is repeated for each sample or groups scanned. With ISP-ARS the FTARS data is used as a predictor and an ISP acoustic waveform is constructed from the prediction (E). Once the ISP waveform is constructed the traditional FTARS cycle is not

needed. Samples scanned with the ISP waveform are classified according to their voltages (F-G). If a sample cannot be classified (H) then FTARS is employed and a new ISP acoustic waveform is constructed chemometrically with a training set that it includes the new unknown. As samples change the ISP waveform can adapt to the new data.

FTARS is well established and has been shown to differentiate drugs, powders, liquids, as well as predict dissolution rate in otherwise identical samples. FTARS is nondestructive and complete scans can be made in seconds, therefore it should be a prime candidate for use as a PAT.

Unfortunately, FTARS relies on intensive computer processing following data collection due to the amount of information gained in each scan. An ARS spectrum recorded over the interval of 20 Hz to 20 kHz with a sample rate of 44.1 kHz for one second generates a substantial amount of data (44100 data points). Chemometric analysis of multiple FTARS data sets are computationally demanding and limits the production rate of tablets, especially if 100% tablet inspection is considered.

ISP-ARS reduces the computational burden of FTARS because it directly produces the analyte identity as an output. ISP-ARS is fast enough to not limit production rates.

V. RESULTS

ISP acoustic waveforms composed of 10, 100, and 1000 frequencies were used to identify several toll-manufactured drugs. The pills used in this study were aspirin, acetaminophen, D-tagatose, ibuprofen, vitamin C, and vitamin B. It was found that only the top 10 frequencies were required to properly classify each pill used in this study. Intra-cluster distances were calculated to be less than 3 multidimensional standard deviations (MSD) for each pill type. The average accuracy of prediction was 98.47, 97.45 and 95.41 percent for the 10, 100 and 1000 frequency component acoustic waveforms respectively [7], [8].

The study was performed in a laboratory under controlled circumstances using prototype ISP_ARS devices custom built for the purpose. The results can be improved significantly by having the devices constructed by a professional sensor building company. For practical use of the ISP-ARS of this paper on a manufacturing line or in a hospital or clinic, this option is clearly necessary and viable financially.

VI. CONCLUSIONS AND FUTURE PLANS

We have described a prototype DDDAS for use in identifying defective or mislabeled pills. Integrated sensing processing acoustic resonance spectroscopy has the ability to differentiate between different types of pills in contract manufacturing and bedside applications.

The results are preliminary and much more research and development will be necessary in order to produce systems that can be deployed on pharmaceutical manufacturing lines. A large number of pills need to be analyzed to build comprehensive libraries. We have been building libraries using the simple technique that whenever any member of the research team is prescribed a medication not in the library, we analyze it and add it. This is nondestructive procedure so the pill is still useful.

We need to design and produce production quality devices for two quite different environments: (1) for manufacturing lines in which pills go by very quickly with with considerable background noise, and (2) a hospital or clinic so that caregivers can correctly identify all pills before giving them to patients.

In the first environment, the sensors will be in fixed locations as part of a government approved production facility. The major obstacle is implementing an enclose that reduces the background noise to a level that can be easily filtered out by the sensor.

In the second environment, a handheld version must be networked wirelessly. There should be little background noise and the device has to be sturdy enough to survive inevitable drops and daily wear and tear from caregivers.

In both cases, substantial computing resources will be needed on occasion to create or update libraries.

REFERENCES

[1] Linda T. Kohn, Janet M. Corrigan, and Molla S. Donaldson (eds.), “To Err is Human”, Journal, National Academies Press, Washington, 2000.

[2] 403rd Wing Public Affairs, Novermber 4, 2007. Available: http://www.403wg.afrc.af.mil/news/story.asp?id=123066690. Site visited 12/29/2008.

[3] U.S. Food and Drug Administration, September 3, 2007. Available: http://www.fda.gov/bbs/topics/NEWS/2006/NEW01507.html. Site visited 11/25/2007.

[4] Woodcock, Janet. US Food and Drug Administration. Available: http://www.fda.gov/ohrms/dockets/ac/02/briefing/3869B1_08_woodcock/sld001.htm. Site visited 11/25/2007.

[5] M. Parashar, V. Matossian, H. Klie, S.G. Thomas, M. Wheeler, T. Kurc, J. Saltz, R. Versteeg, Towards Dynamic Data-Driven Management of the Ruby Gulch Waste Repository, in Computational Science - ICCS 2006: 6th International Conference, Reading, UK, May 28-31, 2006, Proceedings, Part III, Vassil N. Alexandrov, Geert Dick van Albada, Peter M.A. Sloot, Jack J. Dongarra (eds.), Lecture Notes in Computer Science series, vol. 3993, Springer-Verlag Heidelberg, 2006, pp. 384-392.

[6] National Science Foundation January 2006 DDDAS Workshop Report. Available: http://www.dddas.org/nsf-workshop-2006/wkshp_report.pdf. Site visited 12/29/2008.

[7] J. P. Medendorp, J. A. Fackler, C. C. Douglas, and R. A. Lodder, Integrated sensing and processing acoustic resonance spectrometry (ISP-ARS) for sample Classification, J. Pharm. Innov, 2 (2007), pp. 125-134.

[8] T. Hannel, D. Link, and R. A. Lodder, ISP-ARS in discriminating Tagatose and other toll manufactured drugs, ISPE Journal of Pharmaceutical Innovation, 3 (2008), pp. 152-160.