Journal International of Modern Engineering
WWW.IJME.US
Print ISSN: 2157-8052
Online ISSN: 1930-6628

INTERNATIONAL JOURNAL OF MODERN ENGINEERING

Fall / Winter 2011
Volume 12, No. 1

The Leading Journal of Engineering, Applied Science and Technology

Industrial

Electronics

Biomedical

Civil

Aerospace

Computer

Electrical

Chemical

Mechanical

ENGINEERING

Time to change the way academics works? Try acamedics!

If you are not using the Internet, please skip this ad and enjoy the rest of the journal. The average burden of reading this ad is two minutes, which may add two months of productivity to your academic life. Seriously!

Whether you organize a conference, publish a journal, or serve on a committee to collect and review applications, you can use the Internet to make your life easier. We are not talking about emails. We are talking about state-of-the-art online systems to collect submissions, assign them to reviewers, and finally make a decision about each submission. We are talking about value-added services, such as payment and registration systems to collect registration fees online. We are talking about digital document services, such as proceedings CD/DVD development and duplication, or creating professional-looking digital documents.

Finally, we are talking about AFFORDABLE PRICE, QUALITY, and CUSTOMIZATION. And we are talking about each of them at the same time.

By the way, you don't have to be a computer geek to use our systems. We have a couple of them, and they will do all the technical mumbo jumbo for you. We also have a few select people from academics like you, and they know what you do. You just relax and enjoy our systems.

If you are still reading this ad, chances are you are interested in our systems or services. So, visit us at *www.acamedics.com*. While you are there, check the names of our clients as well. Most of them are quite familiar, but the list is too long to include here.

ACAMEDICS
Enjoy Technology!

Acamedics.com • 44 Strawberry Hill Ave Suite 7B, Stamford, CT 06932 • Phone: 203.554.4748 • sales@acamedics.com •

INTERNATIONAL JOURNAL OF MODERN ENGINEERING

The INTERNATIONAL JOURNAL OF MODERN ENGINEERING (IJME) is an independent, not-for-profit publication, which aims to provide the engineering community with a resource and forum for scholarly expression and reflection.

IJME is published twice annually (Fall and Spring issues) and includes peer-reviewed articles, book and software reviews, editorials, and commentary that contribute to our understanding of the issues, problems, and research associated with engineering and related fields. The journal encourages the submission of manuscripts from private, public, and academic sectors. The views expressed are those of the authors and do not necessarily reflect the opinions of IJME or its editors.

EDITORIAL OFFICE:

Mark Rajai, Ph.D.
Editor-in-Chief
Office: (818) 677-2167
Email: ijmeeditor@iajc.org
Dept. of Manufacturing Systems Engineering & Management
California State University-Northridge
18111Nordhoff Street
Northridge, CA 91330-8332

THE INTERNATIONAL JOURNAL OF MODERN ENGINEERING EDITORS

TABLE OF CONTENTS

EDITOR'S NOTE:
AIRCRAFT AND THE ART OF FLYING MACHINES:
VTOL STYLE

Philip Weinsier, IJME Manuscript Editor

Lest you think that your time spent reading this issue will be limited to but one field of engineering, we have included a broad range of articles and are certain that you will find something to pique your curiosity. Robotics: Multilink robots are widely used in the manufacturing industry, and the motion control issues of these robot systems have become popular research topics since the first appearance of the robots in industry (p.10). Computers: Two studies dealing with overlay multicast service using P2P connections (p.18) and large-scale, high-volume, stored-video streaming service over P2P networks (p.37). Industry: Fatigue failure represents a significant portion—80-90%—of failure problems in mechanical and structural systems (62). Electronics: The authors of this study looked at the problem of gate leakage current of MOS devices. The reduction of the gate dielectric thickness is one of the core reasons for increases in current due to direct tunneling of electrons through the SiO_2 (p.71).

Few topics, though, have captured the fascination of audiences like that of aircraft and the art of flying machines, myself included. Thus, I was captivated by this study of quadrotor aerial vehicles in which the authors outline their work on implementing a lightweight, onboard controller (p.5). When talking about sophisticated systems requiring at least four control loops—one each for roll, pitch, yaw and altitude—many controls methods end up requiring large, powerful and computationally intense processors, a fact that functionally limits small, mass-sensitive vehicles. And while there have been several successful autonomous flights of quadrotor vehicles, they have all been base-station controlled.

Traditional, fixed-wing aircraft have for decades served the consumer transport industry quite well as they are used in wide open spaces and have long runways for takeoffs and landings. They are not, then, suitable for maneuvering in tight spaces, hovering or taking off and landing vertically (VTOL), though many attempts have been made over the years to achieve such capabilities. In most respects, helicopters offer virtually all of these features, but the advantages proffered by their long rotor blades are offset by a top speed of about 250 mph, control issues related to torque, massive mechanical linkages for varying the pitch angle of the blades, heavy maintenance schedules and high in-flight kinetic energy resulting in large-scale damage should the blades hit anything, all of which translates into high risk for the machine's surroundings.

Early on, quadrotor designs were thought to be potential solutions to problems associated with vertical flight. Early attempts at VTOL technology date as far back as the 1930s with the introduction of a beast called a cyclogyro that included cylcoidal rotors whose blades rotated around a horizontal axis creating both lift and thrust. A decade or two later, the U.S. and other countries went on the hunt for functional VTOL vehicles, an era that brought us designs such as the XFY-1 Pogo, the French *Coléoptère,* the Vertijet, an Airbike, an Aircar from the United Kingdom, the VZ-9V Avrocar (*Avro Canada's* flying saucer) and others too numerous to list here. Current U.S. VTOL technology has yielded two major military designs: the Bell Boeing V-22 Osprey—virtually a 50-50 blend of helicopter and airplane using a tiltrotor design—and the Harrier family of jets using directed jet thrust.

To be sure, we could devote many issues to this topic, so I will stifle my enthusiasm for the moment and allow you to read on.

Editorial Review Board Members

Listed here are the members of the IAJC International Review Board, who devoted countless hours to the review of the many manuscripts that were submitted for publication. Manuscript reviews require insight into the content, technical expertise related to the subject matter, and a professional background in statistical tools and measures. Furthermore, revised manuscripts typically are returned to the same reviewers for a second review, as they already have an intimate knowledge of the work. So I would like to take this opportunity to thank all of the members of the review board.

As we continually strive to improve upon our conferences, we are seeking dedicated individuals to join us on the planning committee for the next conference—tentatively scheduled for 2013. Please watch for updates on our website (**www.IAJC.org**) and contact us anytime with comments, concerns or suggestions. Again, on behalf of the 2011 IAJC-ASEE conference committee and IAJC Board of Directors, we thank all of you who participated in this great conference and hope you will consider submitting papers in one or more areas of engineering and related technologies for future IAJC conferences.

If you are interested in becoming a member of the IAJC International Review Board, send me (Philip Weinsier, IAJC/IRB Chair, philipw@bgsu.edu) an email to that effect. Review Board members review manuscripts in their areas of expertise for all three of our IAJC journals—IJME (the International Journal of Modern Engineering), IJERI (the International Journal of Engineering Research and Innovation) and TIIJ (the Technology Interface International Journal)—as well as papers submitted to the IAJC conferences.

Kevin Berisso	Ohio University (OH)
John Burningham	Clay State University (GA)
Isaac Chang	Cal Poly State University (CA)
Rigoberto Chinchilla	Eastern Illinois University (IL)
Michael Coffman	Southern Illinois Uni-Carbondale (IL)
Kanchan Das	East Carolina University (NC)
Z. T. Deng	Appalachian State University (NC)
David Domermuth	Oregon Institute of Technology (OR)
Marilyn Dyrud	Tennessee Tech University (TN)
Ahmed Elsawy	University Technology (MALAYSIA)
Morteza Firouzi	Drexel University (PA)
Vladimir Genis	Bhagwan Parshuram (INDIA)
Nitish Gupta	Indiana University (IN)
Hamed Hakimzadeh	Penn State University (PA)
Bernd Haupt	Missouri State University (MO)
Rita Hawkins	Safety Engineer, Sonelgaz (ALGERIA)
Youcef Himri	Norfolk State University (VA)
Shelton Houston	Indiana University Purdue (IN)
Charles Hunt	Bloomsburg University (PA)
Pete Hylton	Michigan Technological Univ. (MI)
Ghassan Ibrahim	Bowling Green State University (OH)
John Irwin	Ball State University (IN)
Sudershan Jetley	Acadiaoptronics (MD)
Rex Kanu	Alcorn State University (MS)
Khurram Kazi	Uttar Pradesh Techn. Uni. (INDIA)
Ognjen Kuljaca	Penn State University (PA)
Chakresh Kumar	Excelsior College (NY)
Zaki Kuruppalil	Penn State University Berks (PA)
Ronald Land	Central Michigan University (MI)
Jane LeClair	Florida A&M University (FL)
Shiyoung Lee	Eastern Carolina University (NC)
Soo-Yen Lee	Penn State University (PA)
Chao Li	
Jimmy Linn	
Dale Litwhiler	
Mani Manivannan	Eastern Illinois University (IL)
David Melton	Excelsior College (NY)
Sam Mryyan	California State U. Fresno (CA)
Arun Nambiar	Universiti Teknologi (MALAYSIA)
Hamed Niroumand	University of Central Missouri (MO)
Troy Ollison	Central Connecticut State U. (CT)
Basile Panoutsopoulous	Purdue University (IN)
Karl Perusich	Virginia State University (VA)
Thongchai Phairoh	Honeywell International, Inc.
Huyu Qu	Arizona State University (AZ)
John Rajadas	NYC College of Technology (NY)
Mohammad Razani	Arizona State University-Poly (AZ)
Sangram Redkar	Wireless Systems Engineer
Marla Rogers	Baker College (MI)
Anca Sala	Zagros Oil and Gas Co. (IRAN)
Mehdi Shabaninejad	Virginia State University (VA)
Ehsan Sheybani	Bostan Abad Islamic Azad U. (IRAN)
Jalal Taheri	Purdue University North Central (IN)
Li Tan	North Carolina A&T State Uni. (NC)
Li-Shiang Tsay	Georgia Southern University (GA)
Phillip Waldrop	Nanjing U. of Science & Tech (CHINA)
Liangmo Wang	Penn State University (PA)
Tom Warms	Ball State University (ILN)
Baijian (Justin) Yang	Sam Houston State University (TX)
Faruk Yildiz	Uni. Of Maryland Eastern Shore (MD)
Emin Yilmaz	Morehead State University (KY)
Yuqiu You	Jackson State University (MS)
Pao-Chiang Yuan	Missouri Western State Uni. (MO)
Jinwen Zhu	

EMBEDDED ONBOARD CONTROL OF A QUADROTOR AERIAL VEHICLE

Cory J. Bryan, Mitchel R. Grenwalt, Adam W. Stienecker, Ohio Northern University

Abstract

The quadrotor aerial vehicle is a structure that has recently been investigated by several teams due to its inherent ability to hover in place while carrying small payloads. This structure has required significant processing power to adequately control due to the requirement of a minimum of four control loops for stable flight. A pilot-in-the-loop system has been implemented by multiple research teams where control is handled by onboard processing systems such that the onboard system provides the stability control while the pilot gives flight commands via a standard hobby RC system [1]. While several teams around the world have successfully created autonomous flight using the quadrotor platform, they have all relied on significant base-station computing for control of the flight [1] and traditional control methods not meant for changing systems. This is due, in part, to the fact that many control methods are computationally intense and, for a small, mass sensitive vehicle, large and powerful processors are out of the question. For a system requiring at least four control loops (roll, pitch, yaw and altitude), it is a stretch to consider onboard computing for this platform. However, if the autonomous quadrotor is to be of any use in completing a real mission, its computing power will need to be onboard. Fortunately, multi-core microcontrollers have recently come into use for complex systems such as the quadrotor. This paper outlines an attempt on the part of the authors to implement a lightweight, onboard controller for the quadrotor platform, and describes an attempt at adaptively controlling the structure. This controller consists of four multi-core microcontrollers for a total of thirty-two single cores. The system and its control method are described here along with the physical implementation and results.

Introduction

Because the traditional fixed-wing airplane suffers from low maneuverability in tight spaces and does not have the ability to hover in place, micro-UAVs of non-traditional structure have been given greater attention in recent years. One of these designs is the quadrotor. The quadrotor is a helicopter design with four dual-axis, symmetrically located points of thrust. This thrust is provided by fixed-pitch rotors, which are commonly driven by four electric motors.

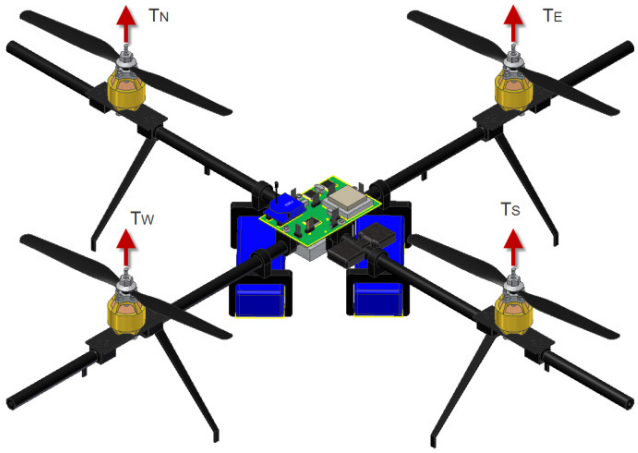


Figure 1. The Quadrotor Aerial Structure

Normally, the flight control electronics are mounted such that the center of gravity falls into the center of the structure at a level slightly below the line of thrust for added stability [1], [2]. A traditional helicopter possesses a tail with a rotor that is used to counteract the yaw thrust created by the spinning rotor. However, unlike a traditional helicopter, the quadrotor has two rotors spinning clockwise and two rotors spinning counterclockwise to eliminate the yaw thrust, thereby eliminating the need for a tail.

With the use of four fixed-pitch rotors, control is seemingly easier than that of a traditional helicopter with one point of thrust. To pitch or roll the quadrotor, the control system must only unbalance the thrust of two rotors on the same axis. Given that the controlled axes (roll, pitch and yaw) are not singularly related to the controlled motor speeds, a set of transformation equations must be developed. From observation, the following equations (1-4) that describe the action of the structure can be written [3].

$$T_T = T_N + T_S + T_E + T_W \tag{1}$$

$$T_R = l(T_W - T_E) \tag{2}$$

$$T_P = l(T_N - T_S) \tag{3}$$

$$T_Y = \frac{\kappa}{\alpha}(T_N + T_S - T_W - T_E) \tag{4}$$

In Equations (1) through (4), T_T , T_R , T_P , T_Y , T_N, T_E, T_S and T_W represent total thrust, roll thrust, pitch thrust, yaw thrust and motor/rotor thrust from the north, east, south and west motor/rotor combinations, respectively. The constants *l,* k and a represent the length from the center of the structure to the thrust point, the drag coefficient and the thrust coefficient, respectively. Together, these four equations can be represented as a transformation matrix, M, shown in Equation (5) that relates attitude thrust to motor thrust.

$$T_A = MT_M \rightarrow \underbrace{\begin{bmatrix} T_T \\ T_R \\ T_P \\ T_Y \end{bmatrix}}_{T_A} = \underbrace{\begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 0 & -1 & 1 \\ 1 & -1 & 0 & 0 \\ \kappa/\alpha & \kappa/\alpha & -\kappa/\alpha & -\kappa/\alpha \end{bmatrix}}_{M} \underbrace{\begin{bmatrix} T_N \\ T_S \\ T_E \\ T_W \end{bmatrix}}_{T_M} \quad (5)$$

However, because the attitude is being sensed and the rotor thrust controlled, the inverse of the transformation matrix, as given I Equation (6), is more useful.

$$T_M = M^{-1}T_A \Rightarrow \begin{bmatrix} T_N \\ T_S \\ T_E \\ T_W \end{bmatrix} = \begin{bmatrix} 0.25 & 0 & \frac{1}{2l} & \frac{\alpha}{4\kappa} \\ 0.25 & 0 & \frac{-1}{2l} & \frac{\alpha}{4\kappa} \\ 0.25 & \frac{-1}{2l} & 0 & \frac{-\alpha}{4\kappa} \\ 0.25 & \frac{1}{2l} & 0 & \frac{-\alpha}{4\kappa} \end{bmatrix} \begin{bmatrix} T_T \\ T_R \\ T_P \\ T_Y \end{bmatrix} \quad (6)$$

A substantial amount of research by others has yielded a plethora of information on the subject [4]. However, there is one very big difference between these studies and the quadrotor design discussed here. Most quadrotors have the computations and control systems running on a base station, such as a workstation computer, and only contain onboard electronics to sense and perform the calculated control signal received via some wireless link [5], [6]. Realistically, due to many factors including cost and security, the reduction of dependence on a wireless control signal is necessary if these devices are to function long term and in the real world. The quadrotor described in this paper is completely independent of a base station. All sensing, control and actuation needed for flight is done onboard via four multi-core microcontrollers containing a total of thirty-two computing units. While the system is not dependent on a base station, it does send signals back to a monitoring station so that critical information can be monitored during the development phase of this research.

When controlling a flying object, adequately descriptive models are not exactly linear nor are their parameters predictable. The literature provides examples of adequately descriptive models but the models are linearized around an operating point and the design continues from there [3], [7]. These same sources indicate success in flight with near constant external parameters such as temperature and pressure, but no analysis has been found regarding flight performance over a wide operating range that is indicative of the real world. Since the field of adaptive control was created for flight [8], it seems the quadrotor aerial vehicle would benefit from an application of adaptive control given that the model for flight changes with temperature, humidity, pressure and altitude. In this paper, the hardware design of the system that gives the structure onboard control, the methods used for control of the system and preliminary results of the research are discussed.

The Design

To perform adequate control and monitoring of performance, sensing on the quadrotor currently includes GPS, an Attitude Heading Reference System (AHRS), a sonar sensor, temperature sensors and an analog-to-digital converter (ADC) for reading the battery voltage. The AHRS is the "9 DOF Razor IMU" from SparkFun Electronics. Sonar, the Parallax Ping Sensor, is used to measure altitude in low-level operations such as takeoff, landing and low-level flight, where GPS is not accurate enough. Critical to monitoring and the safety of the equipment and researchers are the temperature sensors. Due to the use of lithium-polymer batteries (4 cell, 6Ah), there is a risk of serious damage if the batteries do not stay within an operational voltage range. Therefore, the battery pack temperature and voltage are carefully monitored and displayed on the base station at all times. If any of the variables reach predefined boundaries, they will trigger fail-safe protocols to safely land the quadrotor to prevent damage to itself and others.

Another aspect of monitoring for safety purposes is the ADC used to monitor the battery voltage. This is vital because if the lithium-polymer battery drops below ~13 volts, it could internally damage the battery and cause it to explode in flight or shortly after landing, thereby causing serious damage to the quadrotor or the research team. The quadrotor is equipped with a GPS module (Parallax PMB-248) affixed to the top of the circuit board. GPS is used in real time to give the current latitude, longitude, and an approximate altitude at an update rate of 1Hz. For the purposes of equipment recovery and flight data logging, the

GPS coordinates will also be transmitted back to the monitoring station. For communication between the quadrotor and the base station monitoring station, Zigbee communication modules are used. These modules offer a very reliable, short-range 2.4GHz communication channel ideal for this application. At the other end of the Zigbee transmission is the monitoring center. This system exists to display critical flight information for development purposes; however, near the end of the development phase of this research, the monitoring station will become unnecessary. On the quadrotor, there are four motors with four corresponding motor controllers. The motor controllers are Turnigy TR_B25A for outrunner brushless motors. The motors are Rimfire 35-36-1200kv brushless motors. While the motor controllers connect directly to the battery pack, the onboard electronics require both 3.3V and 5V. For purposes of efficiency, switching mode power supply (SMPS) modules are used to regulate the battery voltage down to the necessary operating voltage.

Since there is a tremendous amount of computing required to maintain stable flight and to stay on course, there are four separate Parallax Propeller multi-core microcontrollers onboard the quadrotor. The microcontrollers can communicate and pass data between themselves so that the computing loads can be separated and balanced. Each of the four microcontrollers has a specific task. The first of the microcontrollers computes the data from the AHRS and passes that data on to another microcontroller. This microcontroller has been left largely unloaded such that a Kahlman filtering approach can be implemented to filter the data as the large motor currents induce significant noise into the sensed AHRS data. The second microcontroller connects to the four motor controllers and is responsible for sending the commanded speed for each motor via a traditional repeating 10-20ms servo pulse. The third microcontroller receives and interprets the GPS information, calculates heading and waypoints, and monitors the temperature sensors because both the GPS and temperature sensors have low bandwidth requirements. The fourth controller is the "central hub" for communication as it routes the information coming from the other three microcontrollers to the appropriate locations, as well as reading the data from the Sonar sensor and broadcasting the monitored information out through the Zigbee module.

When it came time to choose rotors for the quadrotor, few choices were available due to the need for matched tractor and pushers rotors. A tractor rotor creates thrust with a clockwise rotation, whereas a pusher rotor creates its thrust from a counterclockwise rotation. Because of the nature of research and development, the initial choice has been to use inexpensive models, given the initial abuse they have received in the testing phase. The rotors are ten inches in diameter with a pitch of 4.5. This is a very aggressive pitch and provides an excellent amount of thrust. The downside to these rotors is that they are quite flexible and fragile, which results in chips in the leading edge of the rotor and flexing at high speeds, reducing efficiency.

Mechanically, most of the structure is carbon fiber and those points that experience significant force are aluminum. which include the center cross connection piece, the motor mounts, and the motor mount to carbon fiber connectors. On the outside of the motor mounts, a connection is available for a safety ring and a tether for test flights. This connector is made from ABS plastic as a sacrificial break point to prevent a catastrophic event requiring an entire airframe rebuild. Altogether, the system weight is approximately two kilograms. Figure 2 is a picture of the current prototype.

Figure 2. Quadrotor with Onboard Control

Embedded Control

Stable flight control of a quadrotor includes four control loops: pitch, roll, yaw and altitude. The pitch and roll controllers can be identical or near identical, as the model is the same, while the altitude and yaw controllers are fairly simple. The focus of this study was on the roll and pitch controllers. Many other researchers [3], [7] begin with a system model, a linearization attempt around some operating point, a standard control system design, and then use some method to create a digital implementation of the controller. For example, Pounds et al. [3] developed the linearized model around a thrust, T_0, as shown in Equation (7).

$$\dot{T}_i = -AT_i + Bu_i \tag{7}$$

$$where\ i = N, S, E, and\ W,$$

$$A = \frac{2}{\tau} + \frac{3\kappa}{2\sqrt{\alpha}}\sqrt{T_0},\ \ and\ \ B = \frac{2\sqrt{\alpha}\kappa}{\tau}\sqrt{T_0}$$

Their model was used to design a double lead compensator for control in the roll and pitch channels. They went on to say that "suitable control performance was obtained." Of course, since the model was linearized around T_0, the flight control performance was only optimized at this level of thrust. An adaptive control system does not always require a system model; it can develop a model itself through a system identification routine. The particular adaptive controller selected for the pitch and roll channels was an adaptive PID controller based on the success of others using the non-adaptive version of the PID controller on this structure [8].

The chosen adaptive PID algorithm uses the weighted recursive least squares method of identification [9], [10]. It is designed to control systems that can be adequately represented with a second-order equation [10]. While the pre-linearized model given by Pounds et al. [3] was not second order or linear, the approximation of the system as second order was made in order to satisfy the requirements of the algorithm. The generic system model is given in Equation (8).

$$\frac{y(k)}{u(k)} = \frac{b_1 q^{-1} + b_2 q^{-2}}{1 + a_1 q^{-1} + a_2 q^{-2}} = \frac{B(q^{-1})}{A(q^{-1})} \tag{8}$$

The control works by first adding a stable pole to the closed-loop characteristic equation and using pole shifting as shown in Equation (9).

$$\left(1 + a_1 q^{-1} + a_2 q^{-2}\right) \rightarrow \left(1 + \alpha q^{-1}\right)\left(1 + a_1 \alpha q^{-1} + a_2 \alpha^2 q^{-2}\right) \tag{9}$$

In Equation (9), $0 \le a \le 1$. The control equation for the adaptive PID is given in Equation (10) and the PID structure in Equation (11).

$$u(k) = \frac{P(q^{-1})}{R(q^{-1})} y_r(k) - \frac{G(q^{-1})}{R(q^{-1})} y(k) \tag{10}$$

$$u(k) = \frac{-S_T K_i}{(1 - q^{-1})(1 + r_1 q^{-1})}(y_r(k) - y(k)) \tag{11}$$

$$+ K_p y(k) + \frac{K_d (1 - q^{-1})}{S_T (1 + r_1 q^{-1})} y(k)$$

In Equations (10) and (11), K_i, K_p and K_d are the PID gains and S_T is the sampling time. From these two equations, P, G and R can be defined as follows:

$$R(q^{-1}) = (1 - q^{-1})(1 + r_1 q^{-1}),\ G(q^{-1}) = g_0 + g_1 q^{-1} + g_2 q^{-2},\ P(q^{-1}) = -S_T K_i \tag{12}$$

By combining Equations (8) and (10), the closed-loop system can be written as shown in Equation (13).

$$\frac{y(k)}{y_r(k)} = \frac{B(q^{-1})P(q^{-1})}{A(q^{-1})R(q^{-1}) + B(q^{-1})G(q^{-1})} \tag{13}$$

If the resulting characteristic equation is solved by setting it equal to Equation (9), the coefficients can be matched, resulting in a system of simultaneous equations, as shown in Equation (14).

$$\underbrace{\begin{bmatrix} 1 & b_1 & 0 & 0 \\ a_1 - 1 & b_2 & b_1 & 0 \\ a_2 - a_1 & 0 & b_2 & b_1 \\ -a_2 & 0 & 0 & b_2 \end{bmatrix}}_{J} \underbrace{\begin{bmatrix} r_1 \\ g_0 \\ g_1 \\ g_2 \end{bmatrix}}_{Z} = \underbrace{\begin{bmatrix} 1 - a_1 + \alpha + a_1\alpha \\ a_1 - a_2 + a_1\alpha^2 + a_1\alpha^2 \\ a_2 + a_2\alpha^3 \\ 0 \end{bmatrix}}_{H} \tag{14}$$

In order to solve the system of equations, J is inverted and simplified before being implemented into the program as the inverse of a 4x4 matrix which, while tedious, can be completed in closed form especially since many of the cells are zero. Once Z is determined, the PID gains (see Equation 15) can be solved for using Equation (11).

$$K_i = \frac{-(g_0 + g_1 + g_2)}{S_T},\quad K_p = \frac{g_1 + 2g_2}{1 + r_1},$$
$$K_d = S_T\left[\frac{r_1 g_1 - (1 - r_1) g_2}{1 + r_1}\right] \tag{15}$$

These gain values are then used to update the control signal, u(k). In combination with the weighted recursive least squares method, the adaptive PID algorithm described above was implemented into the Parallax Propeller multicore microcontroller. With the combination of the proprietary Parallax Propeller high-level language (called SPIN) and assembly language for the floating-point mathematics, the controller was able to sustain an algorithm iteration time of less than 50ms, an acceptable time given that the sam-

pling time of the feedback (S_T) was on the order of 100ms. While this control algorithm was successfully implemented in an embedded environment, the results were not in accordance with expectations. The adaptive algorithm and linearized model were verified using Matlab and a suitable performance was achieved. The adaptive PID controller, derived above, is sufficient for systems that can be modeled by a second-order system—see Equation (8). Based on tests, it was determined that the quadrotor structure cannot be sufficiently modeled for the purposes of adaptive control with a second-order system. Had the linearized model more closely approximated the system, the adaptive controller may have been better able to provide suitable control.

Upon completion of the adaptive PID control testing, work began to implement a standard PID controller for the pitch, roll and altitude channels with the intent to implement a gain scheduling adaptive PID controller. As others have found, the traditional digital implementation of the PID controller is suitable for controlling the quadrotor in flight. However, while the channel controllers have been tuned using standard methods for a single thrust point at hover, the team will be continuing to develop adaptivity into the control strategy. Since suitable PID gains have been determined for each of the channels, the strategy currently under development is to fall back to the adaptive PID control explained here and set initial conditions on the calculated PID gains equal to the traditionally determined PID gains found by tuning. Furthermore, the team plans to limit the calculated PID gains to a range above and below the traditionally gathered PID values such that the adaptive control algorithm has freedom to auto-tune the loop but not enough freedom to cause a catastrophic flight pattern.

Conclusion

The hardware design was successfully implemented on a custom designed, two-layer, printed circuit board (PCB). Using surface mount technology whenever possible allowed the creation of a very small PCB measuring 2.5” x 3.75”. With the exception of the batteries, motors, motor controllers, sonar sensor and temperature sensors, all components occupied real estate on the PCB. Care was taken to locate off-board connections nearest to the place to which they would connect to minimize weight in cabling and unnecessary complications. The placement of the motor controllers was also carefully considered as they must dissipate significant heat during heavy current flow. They were located directly below the flow of air under the rotor. While this reduces the lift achieved, it does allow the components to run cooler. Problems were encountered, but eventually the mechanical and electrical components were made fully functional and capable of executing the control developed on board the quadrotor.

For purposes of testing the control system, single-axis tests were used. These tests required a test rig that would balance the quadrotor on one axis so that the control system could be tuned along each axis. This rig consisted of two tripods that were used to hold two mounted plastic bearings to which the quadrotor was mounted. While this cannot assist in the testing of multiple-axis control, it provided a starting point. Once suitable single-axis control was achieved, tethering was used to test the combination of the control in multiple axes along with on-line tuning of the altitude control.

Battery life is a major concern with any UAV. From a physics standpoint alone, it takes more energy to hover than to glide through the air [11]. The use of more motors makes the craft more maneuverable but also requires more power. With the initial motors, a flight time of fifteen to eighteen minutes was calculated depending on flight conditions such as temperature and wind. However, once the motors were upgraded, the flight time was reduced to ten to thirteen minutes.

From a control standpoint, it was determined that the adaptive PID controller based on a second-order model is not sufficient for the control of a quadrotor on the pitch or roll channels due to the insufficiency of the second-order model to adequately model the quadrotor. However, the adequacy of the standard PID control was validated on the pitch, roll and altitude channels. Further work is being considered using the known tuning parameters from the non-adaptive PID in the adaptive PID to both inform the algorithm of a good starting point and to limit the gains to some range around the known tuning parameters.

References

[1] Fowers, S. (2008). *Stabilization and Control of a Quad-rotor Micro-UAV Using Vision Sensors. Master’s Thesis*, Bringham Young University.

[2] Pounds, P., Mahony, R. & Corke, P. (2006). Modeling and Control of a Quad-Rotor Robot. *Proceedings of the Australasian Conference on Robotics and Automation*.

[3] Pounds, P., Mahony, R. & Mahony, R. (2002). *Design* of a Four-Rotor Aerial Robot. *Proceedings of the Australasian Conference on Robotics and Automation*.

[4] Stepaniak, M. (2008). *A Quadrotor Sensor Platform*. PhD Dissertation, Ohio University.

[5] Castillo, P., Dzul, A. & Lozano, R. (2004). Real-Time Stabilization and Tracking of a Four Rotor Mini-Rotorcraft. *IEEE Transactions on Control Systems Technology*, 12(4), 510-516.

[6] Hoffman, G., Rajnarayan, D., Waslander, S., Dostal, D., Jang, J. & Tomlin, C. (2007). *The Stanford Testbed of Autonomous Rotorcraft for Multi Agent Controls (STARMAC)*. Unpublished paper.

[7] Bouabdallah, S., Noth, A. & Siegwart, R. (2004). PID vs LQ Control Techniques Applied to an Indoor Micro Quadrotor. *Proceedings of the International Conference on Intelligent Robots and Systems*, (pp. 2451-2456).

[8] Astrom, K. J. & Wittenmark, B. (2008). *Adaptive Control*. Mineola, NY: Dover Publications.

[9] Ghandakly, A. (2004). *Recursive Least Squares Identification of Dynamic Systems and an Introduction to Self Tuning Regulators*. Adaptive Control. Lecture. University of Toledo, Toledo.

[10] Ghandakly, A. (2006). *Self Tuning PID Controller*. Adaptive Control. Lecture. University of Toledo, Toledo.

[11] Lafleur, J. (2006). Derivation and Application of a Method for First-Order Estimation of Planetary Aeriel Vehicle Power Requirements. *Proceedings of the 4th International Planetary Probe Workshop*.

Biographies

CORY J. BRYAN is currently a senior in the Department of Technological Studies at Ohio Northern University pursuing a B.S. in Manufacturing Technology. He plans to obtain employment upon graduation at a location where he is able to carry out his interests developed during this research. Dr. Bryan can be reached at c-bryan.1@onu.edu

MITCHEL R. GRENWALT is currently a senior in the Department of Technological Studies at Ohio Northern University pursuing a B.S. in Manufacturing Technology. He has obtained employment and will begin upon graduation. Dr. Grenwalt can be reached at m-grenwalt@onu.edu

ADAM W. STIENECKER teaches electronics and applied control systems courses at Ohio Northern University in the Department of Technological Studies. He holds undergraduate and doctorate degrees in Electrical Engineering from the University of Toledo in Ohio. His areas of research include embedded systems and advanced control of mobile robots. Dr. Stienecker can be reached at adam@drstienecker.com

FRACTIONAL ORDER PID DESIGN FOR ROBOTIC NONLINEAR MOTION CONTROL

Yuequan Wan, Haiyan Henry Zhang, Richard Mark French, Purdue University

Abstract

The modeling of multilink robots produces typical nonlinear systems with uncertain disturbances and high-order matrices. The authors present a method of applying a fractional -order PID controller to such a nonlinear system and show the advantages of this fractional controller. In this study, the dynamic model of the system served as the foundation to derive the control law and objective function for the optimization design of the subjected fractional-order control system. The frequency domain closed-loop transaction function of this fractional system was developed and is discussed here along with controllability, observability and robust satiability. The authors demonstrated the use of algorithms to design and optimize the fractional-order PID to the nonlinear motion control system. By conducting a series of numerical computations, the authors showed that the fractional-order PID controller could enlarge the stable region of a multilink robot system and, therefore, deliver superior control performance in terms of trajectory tracking. The results and procedures introduced here could be practically generalized to other similar systems.

Introduction

Multilink robots are widely used in the manufacturing industry, and the motion control issues of these robot systems have became popular research topics for decades since the first appearance of the robots in industry. Generally speaking, multilink robot systems typically are nonlinear and always involve disturbances. The fine control of industrial robots usually requires complex control systems, careful calibrations and optimizations. In practice, most of these multilink robots are controlled by PID controllers which have the merits or effectiveness, simplicity and feasibility. Although ordinary PID controllers can achieve satisfactory results in most common manufacturing missions, they still lack enough precision in the field and often require precise instrument control.

The ordinary PID controller is designed to provide the restoring, corrective and counteractive forces to the controlled system. In typical situations, the ordinary PID controller can always effectively achieve the control objectives without obvious drawbacks. However, in modern industry, the demand for precise control is driving people to search for improvements. Fractional-order PID (FoPID) introduced here is a natural extension to ordinary PID controllers based on the fractional calculus theory. Since in fractional calculus the orders of integral and derivative are not limited to integer orders anymore, a new type of PID controller can be introduced by replacing the ordinary order integrators and differentiators with fractional-order ones. The main advantages of the FoPID controllers include an enlarged stable region, relatively feasible structure and raised control precision.

As mentioned above, fractional calculus takes the order of integrals and derivatives as any real number. It has a history nearly as long as ordinary calculus, which considers only integer orders [1]. Recently, successful applications of this technology have been found in many fields, such as viscoelasticity [2], [3], control theory [4], [5] and electroanalytical chemistry [6], [7]. In control theory, the general conclusion about a fractional control system is that it could enlarge the stable region [8] and yield a performance at least as good as its integer counterpart. Another important advantage is that fractional integrals or derivatives are hereditary functional while the ordinary ones are point functional. It is known that the hereditary function has a long memory characteristic [9], which means that at any time it would process a total memory of past states. This unique characteristic serves as one of the important reasons for its better performance. For FoPID controllers, many scholars have made tremendous contributions in recent years [10], especially in the tuning rules [11], [12], approximation [13] and stability conditions [14]. These previous studies drove the foundation for the work done in this study.

In this study, then, the authors applied FoPID controllers to a nonlinear multilink robot system and take uncertain disturbances into consideration. Furthermore, the fractional orders of the integrators and differentiators used here are considered as design variables rather than pre-defined parameters. The authors studied the stability conditions and optimization design method for the overall comprehensive performance of the FoPID controllers on the basis of the mathematical model of an Adept 550 robot [15]. Adept 550 is widely used in the industry. It has four axes with three rotational joints and one translational joint. Its beauty lies in its small motion envelope, high speeds and payloads. These

features make the Adept 550 robot a feasible tool for fast and precise operations in production lines, such as subassembly and assembly, packaging and even driving screws. The authors' complete study of FoPID controllers using the Adept 550 robot shows that the fractional controller could achieve high precise control and bring feasible approaches to optimize the design of the FoPID in other applications.

Dynamic Model of an Adept 550 Robot

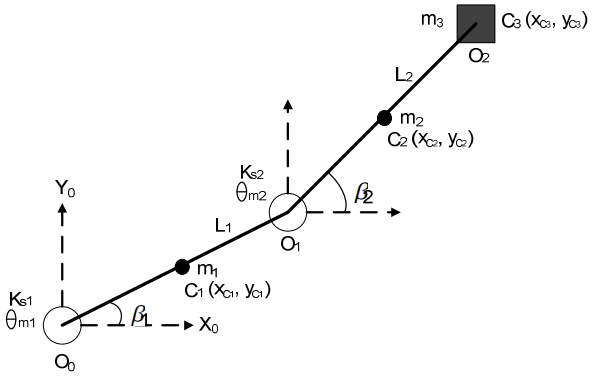


Figure 1. Simplified Structure of Adept 550 Robot

Table 1. D-H Parameters

Link	Li	i	di	i
Inner	L1	0	0	1
Outer	L2	0	0	2

$$A_i = R_{z,\theta_i} T_{z,d_i} T_{x,L_i} R_{x,\alpha_i} = \begin{bmatrix} \cos\theta_i & -\sin\theta_i \cos\alpha_i & \sin\theta_i \sin\alpha_i & L_i \cos\theta_i \\ \sin\theta_i & \cos\theta_i \cos\alpha_i & -\cos\theta_i \sin\alpha_i & L_i \sin\theta_i \\ 0 & \sin\alpha_i & \cos\alpha_i & d_i \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (1)$$

The simplified structure of an Adept 550 robot is shown in Figure 1. When applying Denavit-Hartenberg (D-H) coordinates to Adept 550, one can see the special case of parallel z axes. Its inner and outer links are assumed to be rigid bodies, whose motion determines the trajectory of this robot. Its trajectory is not affected by the gripper angle adjustment during the rotation of the wrist. Without loss of generality, the wrist's rotary angle is assumed to be zero, thus the study could focus on the performance of trajectory tracking. Assuming the notations shown in Figure 1, and using the D-H parameters of the inner (i=1) and outer (i=2) links from Table 1, the following matrix describes the coordinate transformation of rotation and translation:

The subsequent transformation matrix from the base to the gripper can be derived as:

$$T_0^2 = A_1 A_2 = \begin{bmatrix} \cos(\theta_1+\theta_2) & -\sin(\theta_1+\theta_2) & 0 & L_1\cos\theta_1 + L_2\cos(\theta_1+\theta_2) \\ \sin(\theta_1+\theta_2) & \cos(\theta_1+\theta_2) & 0 & L_1\sin\theta_1 + L_2\sin(\theta_1+\theta_2) \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (2)$$

The gripper's horizontal position (P_x, P_y) can be expressed as

$$\begin{aligned} P_x &= L_1 \cos\beta_1 + L_2 \cos\beta_2 \\ P_y &= L_1 \sin\beta_1 + L_2 \sin\beta_2 \end{aligned} \quad (3)$$

where β_i (i = 1,2) is the angular position of the motors, and θ_i (i = 1,2) is the angle about previous z from old x to new x . The relationships of them are described as $\beta_1 = \theta_1$, $\beta_2 = \theta_1 + \theta_2$. Thus, we have the motor's angular positions (β_1, β_2), forward velocity v and backward velocity $(\dot{\beta}_1, \dot{\beta}_2)$, and backward acceleration $(\ddot{\beta}_1, \ddot{\beta}_2)$ [15]:

$$\begin{aligned} \beta_1 &= 2\tan^{-1}\left(\frac{P_y \pm \sqrt{P_x^2 + P_y^2 - R_1}}{P_x + R_1}\right) \\ \beta_2 &= 2\tan^{-1}\left(\frac{P_y \pm \sqrt{P_x^2 + P_y^2 - R_2}}{P_x + R_2}\right) \end{aligned} \quad (4)$$

where

$$\begin{aligned} R_1 &= \frac{P_x^2 + P_y^2 + L_1^2 - L_2^2}{2L_1} \\ R_2 &= \frac{P_x^2 + P_y^2 + L_2^2 - L_1^2}{2L_2} \end{aligned} \quad (5)$$

$$v = \begin{pmatrix} \dot{P}_x \\ \dot{P}_y \end{pmatrix} = J_a \begin{pmatrix} \dot{\beta}_1 \\ \dot{\beta}_2 \end{pmatrix} \quad (6)$$

and where J_a is a Jacobian matrix

$$J_a = \begin{bmatrix} -L_1 \sin\beta_1 & -L_2 \sin\beta_2 \\ L_1 \cos\beta_1 & L_2 \cos\beta_2 \end{bmatrix} \quad (7)$$

$$\begin{pmatrix} \dot{\beta}_1 \\ \dot{\beta}_2 \end{pmatrix} = J^{-1} \begin{pmatrix} \dot{P}_x \\ \dot{P}_y \end{pmatrix} \quad (8)$$

$$\begin{pmatrix} \ddot{P}_x \\ \ddot{P}_y \end{pmatrix} = J_a \begin{pmatrix} \ddot{\beta}_1 \\ \ddot{\beta}_2 \end{pmatrix} + J_v \begin{pmatrix} \dot{\beta}_1^2 \\ \dot{\beta}_2^2 \end{pmatrix} \quad (9)$$

and where

$$J_v = \begin{bmatrix} -L_1 \cos\beta_1 & -L_2 \cos\beta_2 \\ -L_1 \sin\beta_1 & -L_2 \sin\beta_2 \end{bmatrix} \quad (10)$$

$$\begin{pmatrix} \ddot{\beta}_1 \\ \ddot{\beta}_2 \end{pmatrix} = J_a^{-1} \begin{pmatrix} \ddot{P}_x \\ \ddot{P}_y \end{pmatrix} - J_a^{-1} J_v \begin{pmatrix} \dot{\beta}_1^2 \\ \dot{\beta}_2^2 \end{pmatrix} \quad (11)$$

Applying the Lagrange method, the dynamics of the Adept 550 robot can be described [15] as:

$$D(\beta)\ddot{\beta} + H(\beta, \dot{\beta})\dot{\beta} + G(\beta) = \tau + \tau_{damping} \quad (12)$$

where

$$D(\beta) = \begin{bmatrix} \left(\frac{7}{12}m_1 + m_2 + m_3\right)L_1^2 & \left(\frac{1}{2}m_2 + m_3\right)L_1 L_2 \cos(\beta_2 - \beta_1) \\ \left(\frac{1}{2}m_2 + m_3\right)L_1 L_2 \cos(\beta_2 - \beta_1) & \left(\frac{7}{12}m_2 + m_3\right)L_2^2 \end{bmatrix} \quad (13a)$$

$$H(\beta, \dot{\beta}) = \begin{bmatrix} 0 & -\left(\frac{1}{2}m_2 + m_3\right)L_1 L_2 \sin(\beta_2 - \beta_1)\dot{\beta}_2 \\ \left(\frac{1}{2}m_2 + m_3\right)L_1 L_2 \sin(\beta_2 - \beta_1)\dot{\beta}_1 & 0 \end{bmatrix} \quad (13b)$$

$$G(\beta) = \begin{pmatrix} \left(\frac{1}{2}m_1 + m_2 + m_3\right)gL_1 \cos\beta_1 - k_{s1}(r\theta_{m1} - \beta_1) \\ \left(\frac{1}{2}m_2 + m_3\right)gL_2 \cos\beta_2 - k_{s2}(r\theta_{m2} - \beta_2) \end{pmatrix} \quad (13c)$$

$$\tau = \begin{pmatrix} \tau_1 \\ \tau_2 \end{pmatrix} \quad (13d)$$

$$\tau_{damping} = \begin{pmatrix} \tau_{damping1} \\ \tau_{damping2} \end{pmatrix} = \begin{bmatrix} -C_1 & 0 \\ 0 & -C_2 \end{bmatrix} \begin{pmatrix} \dot{\beta}_1 \\ \dot{\beta}_2 \end{pmatrix} = C \begin{pmatrix} \dot{\beta}_1 \\ \dot{\beta}_2 \end{pmatrix} \quad (13e)$$

The damping coefficients are included in matrix C .

Model of Fractional-order PID Controllers

Based on Equation (12), it can be assumed that the motors driving the inner and outer links are of the same type. Dynamics of the two link for k = 1,2 is described as:

$$\begin{aligned} &\sum_{j=1}^{n} d_{jk}(\beta)\ddot{\beta}_j + \sum_{i,j=1}^{n} h_{ijk}(\beta)\dot{\beta}_i\dot{\beta}_j + g_k(\beta) = \tau_k - C_k\dot{\beta}_k, \\ &J_{m,k}\ddot{\theta}_{mk} + \left(B_{m,k} + K_{b,k}\frac{K_{m,k}}{R_k}\right)\dot{\theta}_{mk} = \frac{K_{m,k}}{R_k}V_k - \tau_{m,k} \end{aligned} \quad (14)$$

Since $\beta_k = r\theta_{m,k}$, $\tau_{m,k} = r\tau_k$, where r is the gear ratio, the two dynamic equations of robot link and its driving motor expressed in Equation (14) can be combined into a single equation:

$$J_{eff,k}\ddot{\theta}_{mk} + B_{eff,k}\dot{\theta}_{mk} = KV_k - C_k\theta_{mk} - rd_k \quad (15)$$

Now, for a fractional-order PID controller, $PI^{\lambda}D^{\mu}$, one gets the five design parameters summarized in Table 2.

Table 2. Design Parameters for the Controller

K_P	Coefficient for the proportional term
K_D	Coefficient for the derivative term
K_I	Coefficient for the integral term
μ	Fractional order for the
λ	Fractional order for the

The closed-loop control diagram is shown in Figure 2, while Equation (16) describes the transfer function of this closed-loop system . The fractional derivative used in this study is defined as Caputo's fractional derivative [16].

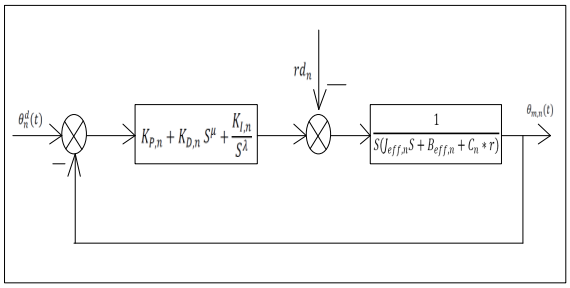


Figure 2. Closed-Loop Diagram of Fractional-Order , $PI^{\lambda}D^{\mu}$ Controlled Robot Arm

$$\theta_{m,n} = \frac{(K_P\theta_n^d - rd_n)S^{\lambda} + K_{I,n}\theta_n^d}{J_{eff,n}S^{\lambda+2} + (B_{eff,n} + C_n r)S^{\lambda+1} + K_{p,n}S^{\lambda} + K_{I,n}} \quad (16)$$

In this study, the FoPID controllers of the two arms had the same fractional order, λ and μ, and different coefficients. Besides, both the fractional order of the integrator and the differentiator are bounded in the range of (0, 1) in this study. In Equation (16), the non-linear terms, d_n , are nonlinear disturbances given as:

$$\begin{aligned} d_1 &= (\tfrac{1}{2}m_2 + m_3)\{L_1 L_2 \cos(\beta_2 - \beta_1)\ddot{\beta}_2 \\ &- L_1 L_2 \sin(\beta_2 - \beta_1)\dot{\beta}_2^2\} + (\tfrac{1}{2}m_1 + m_2 + m_3)gL_1 \cos\beta_1 \\ d_2 &= (\tfrac{1}{2}m_2 + m_3)\{L_1 L_2 \cos(\beta_2 - \beta_1)\ddot{\beta}_1 \\ &- L_1 L_2 \sin(\beta_2 - \beta_1)\dot{\beta}_1^2 + gL_2 \cos\beta_2 \end{aligned} \quad (17)$$

Apply Caputo's fractional-order derivative to Equation (17), and given $\beta_K = r\theta_{m,K}$, the time domain system function could be represented by the following matrix.

$$\begin{bmatrix} J_{eff.1} & r^2T_1 \\ r^2T_1 & J_{eff.2} \end{bmatrix}\begin{bmatrix} {}_0D_t^{\lambda+2}\beta_1 \\ {}_0D_t^{\lambda+2}\beta_2 \end{bmatrix}+\begin{bmatrix} B_{eff.1}+c_1r & r^2T_2\dot{\beta}_2^{2-\lambda} \\ r^2T_2\dot{\beta}_1^{2-\lambda} & B_{eff.2}+c_2r \end{bmatrix}\begin{bmatrix} {}_0D_t^{\lambda+2}\beta_1 \\ {}_0D_t^{\lambda+2}\beta_2 \end{bmatrix}+$$
$$+\begin{bmatrix} K_{P1}+r^2(-T_3\ddot{\beta}_2+T_4\dot{\beta}_2^2+T_5) & r^2(T_3\ddot{\beta}_2-T_4\dot{\beta}_2^2) \\ r^2(-T_3\ddot{\beta}_1+T_4\dot{\beta}_1^2) & K_{P2}+r^2(T_3\ddot{\beta}_1-T_4\dot{\beta}_1^2+T_6) \end{bmatrix}\begin{bmatrix} {}_0D_t^{\lambda}\beta_1 \\ {}_0D_t^{\lambda}\beta_2 \end{bmatrix}$$
$$+\begin{bmatrix} K_{D1} & 0 \\ 0 & K_{D2} \end{bmatrix}\begin{bmatrix} {}_0D_t^{\lambda+\mu}\beta_1 \\ {}_0D_t^{\lambda+\mu}\beta_2 \end{bmatrix}+\begin{bmatrix} K_{I1} & 0 \\ 0 & K_{I2} \end{bmatrix}\begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix}= \quad (18)$$
$$=\begin{bmatrix} K_{P1} & 0 \\ 0 & K_{P2} \end{bmatrix}\begin{bmatrix} {}_0D_t^{\lambda}\beta_1^{\,d} \\ {}_0D_t^{\lambda}\beta_2^{\,d} \end{bmatrix}+\begin{bmatrix} K_{D1} & 0 \\ 0 & K_{D2} \end{bmatrix}\begin{bmatrix} {}_0D_t^{\lambda+\mu}\beta_1^{\,d} \\ {}_0D_t^{\lambda+\mu}\beta_2^{\,d} \end{bmatrix}+\begin{bmatrix} K_{I1} & 0 \\ 0 & K_{I2} \end{bmatrix}\begin{bmatrix} {}_0D_t^{\lambda}\beta_1^{\,d} \\ {}_0D_t^{\lambda}\beta_2^{\,d} \end{bmatrix}$$

where

$$T_1=\left(\frac{1}{2}m_2+m_3\right)L_1L_2\cos(\beta_2-\beta_1) \quad (19a)$$

$$T_2=-\left(\frac{1}{2}m_2+m_3\right)L_1L_2\sin(\beta_2-\beta_1)\frac{\Gamma(3)}{\Gamma(3-\lambda)} \quad (19b)$$

$$T_3=\left(\frac{1}{2}m_2+m_3\right)L_1L_2\cos(\beta_2-\beta_1+\lambda\frac{\pi}{2}) \quad (19c)$$

$$T_4=\left(\frac{1}{2}m_2+m_3\right)L_1L_2\sin(\beta_2-\beta_1+\lambda\frac{\pi}{2}) \quad (19d)$$

$$T_5=\left(\frac{1}{2}m_1+m_2+m_3\right)gL_1\cos(\beta_1+\lambda\frac{\pi}{2}) \quad (19e)$$

$$T_6=\left(\frac{1}{2}m_2+m_3\right)gL_2\cos(\beta_2+\lambda\frac{\pi}{2}) \quad (19f)$$

In Equation (18), the differential order of β_1 β_2 is $\lambda + 2$, $\lambda + 1$, $\lambda + \mu$ and 0. Since these orders are not equally spaced, it is not easy to directly re-write Equation (18) in a linear matrix formation. Inspired by the work of Galkowski et al. [17], it was assumed that λ and μ are rational numbers, which could be expressed by *a/b* and *c/d* and in their relatively prime formats, respectively. By noting that $\beta = [\beta_1 ; \beta_2]$, Equation (18) could be written as:

$$M_{10}D_t^{\frac{ac+2bd}{bd}}\beta+M_{20}D_t^{\frac{ac+bd}{bd}}\beta+M_{30}D_t^{\frac{ac+bc}{bd}}\beta+M_{40}D_t^{\frac{ad}{bd}}\beta+M_{50}D_t^{\frac{0}{b}}\beta-U=0 \quad (20)$$

In Equation (20), M_i and U make up a coefficient matrix with their corresponding terms in Equation (18). One more thing to mention here is that not all of these coefficients are constant, given the uncertain disturbance. Equation (20) actually will be shown later to be a time variant system. By inserting zero matrixes, it is equivalent to rewriting Equation (20) as shown follows:

$$M_{10}D_t^{\frac{ac+2bd}{bd}}\beta+N_{10}D_t^{\frac{ac+2bd-1}{bd}}\beta+N_{20}D_t^{\frac{ac+2bd-2}{bd}}\beta+\cdots+M_{20}D_t^{\frac{ac+bd}{bd}}\beta+$$
$$+N_{i0}D_t^{\frac{ac+bd-1}{bd}}\beta+\cdots+M_{30}D_t^{\frac{ac+bc}{bd}}\beta+\cdots+N_{j0}D_t^{\frac{ad+1}{bd}}\beta+ \quad (21)$$
$$+M_{40}D_t^{\frac{ad}{bd}}\beta+\cdots+N_{k0}D_t^{\frac{a+1}{bd}}\beta+M_{50}D_t^{\frac{0}{b}}\beta-U=0$$

where $N_1 = N_2 = \cdots = N_i = \cdots = N_j = N_K = 0$.

Based on Equation (21) one has an equally spaced fractional -order system on every term and, therefore, the state space could be defined as:

$$x=\begin{bmatrix} {}_0D_t^{\frac{0}{bd}}\beta & {}_0D_t^{\frac{1}{bd}}\beta & \cdots & {}_0D_t^{\frac{ad-1}{bd}}\beta & {}_0D_t^{\frac{ad}{bd}}\beta & \cdots & {}_0D_t^{\frac{ad+2bd-1}{bd}}\beta \end{bmatrix}^T \quad (22)$$

The entire system, then, is:

$${}_0D_t^{\frac{1}{bd}}X=AX+BU \quad (23)$$

where

$$A=\begin{bmatrix} 0_{[2ad+4bd-2,2]} & \vdots & I_{[2ad+4bd-2,2ad+4bd-2]} \\ \cdots & \cdots & \cdots \\ -M_1^{-1}M_5 & 0\cdots -M_1^{-1}M_4 & 0\cdots -M_1^{-1}M_3 & 0\cdots -M_1^{-1}M_2 & 0\cdots \end{bmatrix} \quad (24a)$$

$$B=\begin{bmatrix} 0_{[2ad+4bd-2,2]} & \\ \cdots & \cdots \\ 1 & 0 \\ 0 & 2 \end{bmatrix} \quad (24b)$$

In Equations (24.a) and (24.b), $0_{[2ad+4bd-2,2]}$ is a zero matrix whose dimension is $[2ad + 4bd - 2,2]$ and $I_{[2ad+4bd-2,2ad+4bd-2]}$ is the identity matrix having the dimension of $[2ad + 4bd - 2,2ad + 4bd - 2]$. Equation (23) is the state space representation of our system function. The system matrix A has the dimension of $[2ad + 4bd,2ad + 4bd]$ and *B* has the dimension of $[2ad + 4bd,2]$. The stability study and the design of the fractional-order PID controller will focus on the matrix *A*. Although *A* could have a very high dimension with a different fractional order, the fact that matrix *A* is a sparse matrix makes the task easier in most cases.

Controllable, Observable and Robust Stability of the System

Since matrix *A* is in the controllable canonical form, and consequently one state could be transferred to another, the system is controllable and observable. The design focuses on the robust stability of this system. For a fractional-order system, the system would be guaranteed stable if all of the system matrix's eigenvalues satisfy the following criteria [18].

$$\left|\arg(\lambda)\right| > \beta\frac{\pi}{2} \quad (25)$$

Therefore, in this study, the ratio of the stable region of the FoPID to the integer PID is 2—1/*bd*. One could raise *b* and *d* to get a larger stable region; however, raising them would cause a larger dimension of matrix *A* and involve more eigenvalues since the total number of eigenvalues is *2ad+4bd*. More eigenvalues would make it harder to guarantee that all of them are settled in the stable region.

Moreover, since *A* is a bounded sparse matrix with interval uncertainties, there should be an infinite number of eigenvalues to check to satisfy the stable region if one directly uses the method of Equation (25). In this case, boundaries of each eigenvalue [18], [19] should be checked and the stability of the system—based on the behaviors of all eigenvalue boundaries [20]—continually analyzed. Therefore, the boundaries of this system, matrix *A*, need to be checked. Based on Equations (18) and (19), the following inequality holds:

$$J_{eff,1}J_{eff,2} > r^2 J_m[L_1^2(\frac{1}{2}m_1+m_2+m_3)+L_2^2(\frac{1}{2}m_2+m_3)]+r^4(\frac{1}{2}m_2+m_3)^2L_1^2L_2^2 \quad (26)$$

Thus, the determinant of matrix, M_1, satisfies the condition

$$J_{eff,1}J_{eff,2} - r^4 T_1^2 \neq 0 \quad (27)$$

The fact that the condition in Equation (27) always holds, implies that matrix M_1 is always nonsingular and, consequently, matrix *A* will never be singular if $K_{i1} \neq 0$, $K_{i2} \neq 0$. In this design, the authors kept this condition. Thus, one gets

$$\frac{1}{\det(M_1)} \in [\frac{1}{J_{eff,1}J_{eff,2}}, \frac{1}{J_{eff,1}J_{eff,2} - r^4(0.5m_2+m_3)^2L_1^2L_2^2}] \quad (28)$$

In this robot control study, $\dot{\beta}$ and $\ddot{\beta}$ are also bounded because of reality. Therefore, one should also find that matrix *A* is bounded. Plugging in the parameters used in this study, one gets the following boundary functions for each variant term in *A* through numerical computation, where the boundaries are functions of the design parameters $(K_{I1},K_{I2},K_{P1},K_{P2},K_{D1},K_{D2},\lambda,\mu)$.

$$\overline{-M_1^{-1}M_5} = \begin{bmatrix} -57.2273K_{I1} & 94.6367K_{I2} \\ 94.6367K_{I1} & -229.4504K_{I2} \end{bmatrix}$$
$$\underline{-M_1^{-1}M_5} = \begin{bmatrix} -83.7362K_{I1} & -64.9431K_{I2} \\ -64.9431K_{I1} & -335.73674K_{I2} \end{bmatrix} \quad (29a)$$

$$\overline{-M_1^{-1}M_4} = \begin{bmatrix} 48.1076-71.7260K_{P1} & 27.5032+93.6562K_{P2} \\ 87.6325+71.3310K_{P1} & 98.6555-2623100K_{P2} \end{bmatrix}$$
$$\underline{-M_1^{-1}M_4} = \begin{bmatrix} -44.5903-80.7186K_{P1} & -31.5653-64.9431K_{P2} \\ -96.3535-46.8943K_{P1} & -91.9135-308934.1K_{P2} \end{bmatrix} \quad (29b)$$

$$\overline{-M_1^{-1}M_3} = \begin{bmatrix} -57.2273K_{D1} & 94.6367K_{D2} \\ 94.6367K_{D1} & -229.4504K_{D2} \end{bmatrix}$$
$$\underline{-M_1^{-1}M_3} = \begin{bmatrix} -83.7362K_{D1} & -64.9431K_{D2} \\ -64.9431K_{D1} & -335.73674K_{D2} \end{bmatrix} \quad (29c)$$

$$\overline{-M_1^{-1}M_2} = \begin{bmatrix} -5.7227+0.0711\dot{\beta}_{1,MAX}^{2-\lambda} & 9.4637 \\ 9.4637+0.2069\dot{\beta}_{1,MAX}^{2-\lambda} & -22.945+0.0609\dot{\beta}_{2,MAX}^{2-\lambda} \end{bmatrix}$$
$$\underline{-M_1^{-1}M_2} = \begin{bmatrix} -8.3736 & -6.4943-0.4241\dot{\beta}_{2,MAX}^{2-\lambda} \\ -6.4943-1.7004\dot{\beta}_{1,MAX}^{2-\lambda} & -33.5737 \end{bmatrix} \quad (29d)$$

Now, the robust stability of this FoPID controlled system at different design parameters could be studied $(K_{I1},K_{I2},K_{P1},K_{P2},K_{D1},K_{D2},\lambda,\mu)$. And this feature actually provides a criterion for optimizing the design of the controllers. Next, though, the authors would like to show how the design parameters, which are the coefficients and the fractional order of the two FoPID controllers, affect robust stability. Figure 3 shows this effect. Taking the upper left frame in Figure 3 as an example, the rectangles drawn by blue solid lines show the boundaries of each eigenvalue. Since there are uncertainties involved in this system, the eigenvalues are actually located in a range rather than single spots. And rectangles provide sufficient boundaries for these eigenvalues [19]. To ensure that the system is robustly stable, the eigenvalues' boundaries are not allowed to cross the stable boundary, which essentially is represented by the angle $\pm 2\pi/bd$ in this study. For a better demonstration, the non-violated stable boundaries are plotted by cyan solid lines and those violated stable boundaries by red solid lines. Figure 3 clearly shows that changing the combination of the design variables can change the overall stability of the system. During the design of the entire set of parameters, there could be unlimited permutations for the choices of design variable set $(K_{I1},K_{I2},K_{P1},K_{P2},K_{D1},K_{D2},\lambda,\mu)$. The authors would like to apply some optimization algorithm to achieve the comprehensive optimized design. Since the task of optimization design involves the permutation of each parameter, the genetic algorithm is a natural choice for this mission.

Optimization Design

For this design, the system contains uncertainties and one could only obtain the ranges for each eigenvalue. As shown in Figure 3, the ranges are the rectangles bounded by the four corner eigenvalues. Drawing down these corner eigenvalues in a complex plane and noting their arguments by $\angle\beta_{ij}, i=1,2,\ldots n; j=1,2,3,4$, one could then measure the difference between these arguments and the stable boundary. In this way, and combined with the fact that all eigenvalues are symmetrical to the real axis in the complex plan, a natural optimization objective is to minimize the difference of stable arguments, $2\pi/bd$, to the absolute value of each $\angle\beta_{ij}$

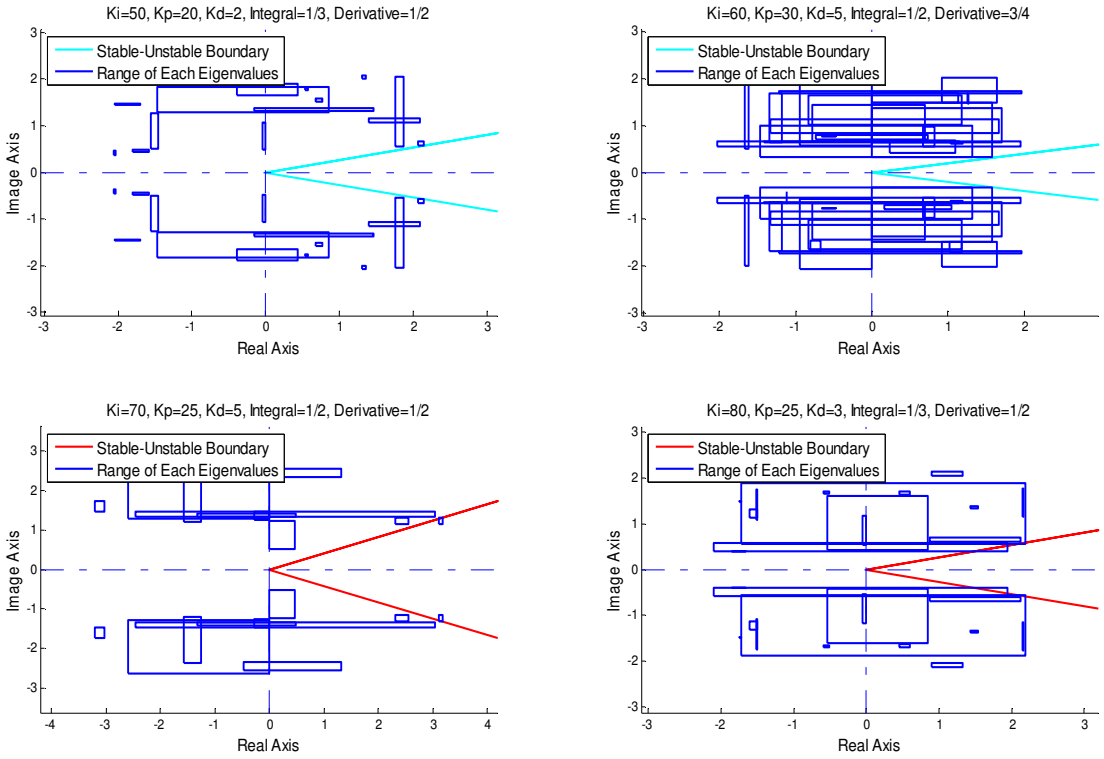


Figure 3. Effect of Changing Design Variables on Overall Stability

Therefore, the optimization function used in this research can be expressed as follows:

$$OptimalDes\ \ ign(K_{I1}, K_{I2}, K_{P1}, K_{P2}, K_{D1}, K_{D2}, \lambda, \mu) = \arg\min \sum_{i=1}^{n} \sum_{j=1}^{4} \psi_{ij} \left(\frac{2\pi}{bd} - \left|\angle\beta_{ij}\right|\right) \quad (30)$$

In Equation (30), ψ_{ij} serves as the coefficient of penalization. There could be many methods used to assign the values of ψ_{ij}, and one could separate the complex plane into different segments according to various criteria. Here, the authors looked at the two-zone and three-zone stepwise penalization methods. Table 3 summarizes these two methods.

Table 3. Value of Penalization Coefficient

Two-Zone Method		Three-Zone Method	
$\left\|\angle\beta_{ij}\right\|$	ψ_{ij}	$\left\|\angle\beta_{ij}\right\|$	ψ_{ij}
$\in [0, 2\pi/bd]$	1e+	$\in [0, 2\pi/bd]$	1e+1
$\in (2\pi/bd, \pi]$	1	$\in (2\pi/bd, 0.5\pi + 2\pi/bd]$	1e+3
N/A	N/A	$\in (0.5\pi + 2\pi/bd, \pi]$	1

Before exploring the trajectory tracking performance, the trajectory planning method used in this study will be introduced. First, let the robot arm move in both the x- and y-directions. Next, set the original point at 500mm by 320mm and allow 1 second for the robot arm to move to position 200mm by 600mm. Figure 4 demonstrates the trajectory plan. Table 4 summarizes the optimization results.

Table 4. Optimization Results of Design Parameters

Two-Zone Method		Three-Zone Method	
K_{P1}	146.93	K_{P1}	93.87
K_{P2}	14.27	K_{P2}	80.60
K_{I1}	67.33	K_{I1}	14.27
K_{I2}	80.60	K_{I2}	146.93
K_{D1}	0.80	K_{D1}	4.70
K_{D2}	3.50	K_{D2}	0.80
λ	0.20	λ	0.67
μ	0.83	μ	0.75

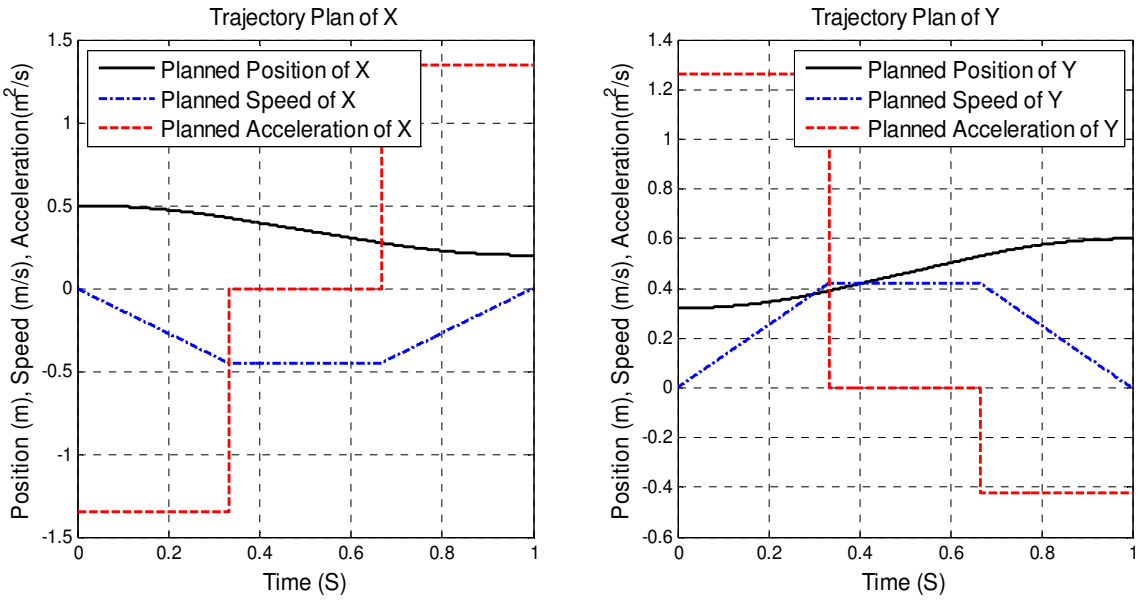


Figure 4. Trajectory Plan

Simulation Results and Conclusions

Figure 5 shows a plot of the simulation results about the trajectory tracking. This figure includes the results from the system optimized by both the two-zone and three-zone methods. And, as a comparison, the authors included the results of an ordinary PID controller [15]. As shown in Figure 5, the optimized FoPID controllers have tracked the trajectory plan successfully. In terms of tracking error, the fractional system achieved a higher precision when compared with the ordinary PID system. Both the two-zone and three-zone methods provided satisfactory optimization results and, therefore, the optimization method studied here can be deemed effective. The tracking error at each sampling point was also recorded and the average squared tracking error computed, as summarized in Table 5. From Table 5, one can clearly see that the FoPID systems have raised the precision of tracking by one order of magnitude.

Table 5. Comparisons in terms of Mean Squared Tracking Errors

	Ordinary PID	FoPID, Two-Zones	FoPID, Three-Zones
Mean Squared Error in X	3.8859e-05	1.2279e-6	7.0355e-6
Mean Squared Error in Y	1.4879e-05	8.8149e-6	3.7683e-6

Evidenced by the simulation results, the FoPID controlled Adept550 robot system could achieve better results in terms of trajectory tracking. And the design methods introduced in this paper are effective for finding the optimized design of the fractional controllers. This method could be easily transferred into other applications related to fractional control and, consequently, bring valuable results to industrial practice.

In summary, then, the following conclusions are offered:

1. The fractional-order control of multilink robot systems always involves disturbance or other uncertainties; therefore, studying the limits of each eigenvalue is a feasible method for evaluating the overall stability. Furthermore, the boundary matrix could be helpful in finding the optimization design of the fractional -order controllers.

2. The stepwise penalized method could be used to optimize the design of FoPID systems, which allows people to move the system's eigenvalues toward to the desired regions. The method proposed in this paper could be generalized to other applications in the design of fractional-order controllers.

3. The optimized fractional system will take advantage of the enlarged stable region, while avoiding any negative effects brought by the increased number of eigenvalues. Simulation results show that the optimized FoPID controlled Adept550 system could track the planned trajectory successfully and raise the

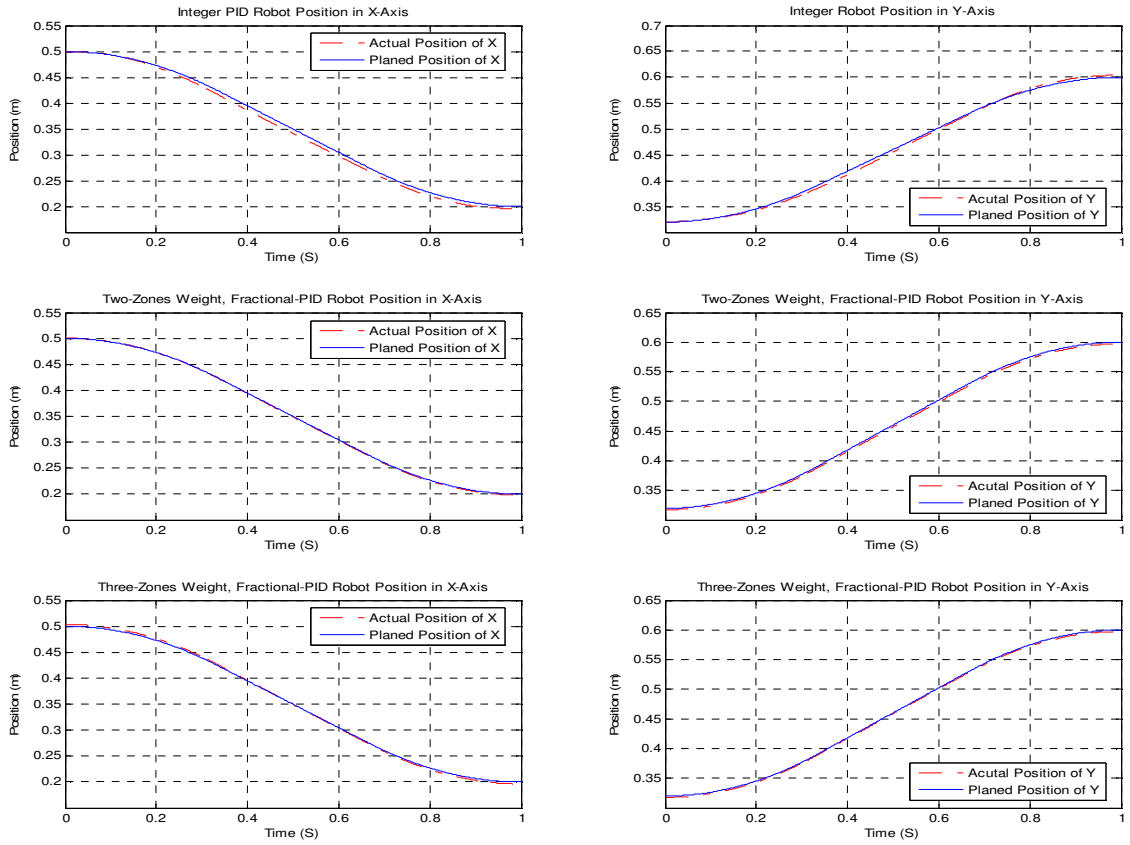


Figure 5. Simulation Comparisons in terms of Trajectory

precision greatly during the tracking process. This characteristic would bring valuable results to the manufacturing industry.

References

[1] Loverro, A. (2004). *Fractional Calculus: History, Definitions and Applications for the Engineer*. Unpublished report. Department of Aerospace and Mechanical Engineering. University of Notre Dame.

[2] Mainardi, F. (1997). *Applications of Fractional Calculus in Mechanics, Transform Methods and Special Functions*. Varna'96, SCT Publishers, Singapore.

[3] Rossikhin, Y. A. & Shitikova, M. V. (1997). Applications of Fractional Calculus to Dynamic Problems of Linear and Nonlinear Hereditary Mechanics of Solids. *Applied Mechanics Reviews*, 50, 15-67.

[4] Bagley, R. L. & Calico, R. A. (1991). Fractional Order State Equations for the Control of Viscoelastically Damped Structures, *Journal of Guidance Control and Dynamics*, 14(2) 304-311.

[5] Makroglou, A., Miller, R. K. & Skkar, S. (1994). Computational Results for a Feedback Control for a Rotating Viscoelastic Beam. *Journal of Guidance, Control and Dynamics*, 17(1), 84-90.

[6] Oldham, K. B. (1976). A Signal Independent Electroanalytical Method. *Analytical Chemistry*, 72, 371-378.

[7] Goto, M. & Ishii, D. (1975). Semi-differential Electro -analysis. *Electroanalytical Chemistry and Interfacial Electrochemistry*, 61, 361-365.

[8] Matignon, D. (1998). *Generalized Fractional Differential and Difference Equations: Stability Properties and Modeling Issues*. Proc. Symposium for Math. Theory of Networks and Systems, Padova, Italy.

[9] Diethelm, K., Ford, N. J., Freed, A. D. & Luchko, Y. (2005). Algorithms for the fractional calculus: A selection of numerical methods. *Computer Methods in Applied Mechanics and Engineering*. 194(6-8), 743-773.

[10] Podlubny, I. (1994). Fractional-Order Systems and Fractional-Order Controllers, Institute For Experimental Physics. *Slovak Academy of Sciences*, 4(2), 28 -34.

[11] Ying, L. & YangQuan, C. (2009). Fractional-order [proportional derivative] controller for robust motion control: Tuning procedure and validation. *American Control Conference*, (pp.1412-1417).

[12] Dingyu, X. & YangQuan, C. (2006). Fractional Order Calculus and Its Applications in Mechatronic System Controls Organizers. *Proceedings of the 2006 IEEE International Conference on Mechatronics and Automation*, (pp. nil33-nil33).

[13] Dingyu, X., Chunna, Z. & YangQuan, C. (2006). A Modified Approximation Method of Fractional Order System. *Proceedings of the 2006 IEEE International Conference on Mechatronics and Automation*, (pp. 1043-1048).

[14] Sabatier, J., Moze, M. & Farges, C. (2010). LMI stability conditions for fractional order systems, Computers & Mathematics with Applications. *Fractional Differentiation and Its Applications*, 59(5), 1594-1609, ISSN 0898-1221.

[15] Zhang, H. (2010). PID Controller Design for A Nonlinear Motion Control Based on Modeling the Dynamics of Adept 550 Robot. *International Journal of Industrial Engineering and Management*, 1(1).

[16] El-Sayed, A. M. A. (1994). Multivalued Fractional Differential Equations. *Applied Mathematics and Computation*, 80, 1-11.

[17] Galkowski, K., Bachelier, O. & Kummert, A. (2006). Fractional Polynomials and nD Systems: A Continuous Case. *Decision and Control, 45th IEEE Conference*, (pp. 2913-2917).

[18] Deif, A. S. (1991). The interval eigenvalue problem. Applied Mathematics and Mechanics, 71(1), 61–64.

[19] YangQuan, C., Hyo-Sung, A. & Podlubny, I. (2005). Robust stability check of fractional order linear time invariant systems with interval uncertainties. *Mechatronics and Automation, 2005 IEEE International Conference*, 1(1), 210-215.

[20] Qiu, Z., Müller, P. C. & Frommer, A. (2001). An approximation method for the standard interval eigenvalue problem of real nonsymmetric interval matrices. *Communications in Numerical Methods in Engineering*, 17, 239-251.

Biographies

YUEQUAN WAN obtained PhD degree from Purdue University in May 2011. His research focuses on applied fractional calculus and fractional order control.

H. HENRY ZHANG is an assistant professor at Purdue University, his research covers control, applied fractional calculus, hydraulics, machining, and multi-discipline design optimization. Dr. Zhang got his PhD from the University of Michigan in June 1996. Dr. Zhang may be reached at hhzhang@purdue.edu

R. MARK FRENCH is an associate professor at Purdue University. his research focuses on applied fractional calculus, music acoustics, aero elastics and optimization. Dr. French got his PhD from the University of Dayton. Dr. French may be reached at rmfrench@purdue.edu

ON THE PERFORMANCE OF AN APPLICATION LAYER MULTICAST PROTOCOL

Xiaobing Hou, Central Connecticut State University; Shuju Wu, Central Connecticut State University

Abstract

In this study, the performance of an application layer multicast protocol, namely Adaptive Overlay Multicast (AOM) protocol was evaluated. In this paper, the authors introduce the concepts of fan-outs and foster limits in building application layer multicast trees that exemplify the performance and adaptability characteristics of AOM and network dynamics with extensive simulation results.

Introduction

Overlay multicast (also called application layer multicast) was proposed to provide multicast service at the application layer using P2P connections, thereby removing the dependence on multicast support of the underlying networks [1-4]. Operations such as membership management and routing are implemented at the application layer, and data distribution is over a multicast tree that consists exclusively of end hosts and unicast connections. Both IP and overlay multicast use a tree to achieve distribution efficiency, which results in performance clusters when tree links or nodes are under fault. A performance cluster includes those members that suffer the performance degradation caused by the same fault. The larger the cluster size, the more the group communication is jeopardized. Therefore, tree construction and fault adaptation are important.

The Internet is dynamic and unpredictable in nature. Dynamic events such as group membership changes, node failures, link failures or network congestion can cause the quality of an overlay multicast tree to degrade over time. Any such events are considered faults. A fault caused by dynamic group membership or node failure is easier to detect than the others and the effect on application performance is temporal. However, faults caused by network congestion in the Internet could last much longer. They also cause end-to-end performance degradation without a total loss of connection and, thus, cannot be detected by simple node failure detection mechanisms. Experiments on the MBone [5] have shown that even for a small multicast group of 11 members, each member experiences a very long consecutive loss of up to a few minutes, a situation that occurs in almost every trace. Link loss rates in an MBone group of eight members are measured in one-hour intervals and have been shown to vary between 2% and 35% [6]. On a specific link, loss rates higher than 15% occur frequently and often last about 10 minutes. Also, from the results of other Internet measurements [7], [8], it is not unusual to find long-lasting high-loss periods between Internet nodes, although the average loss rate over a day could be low. When such faults happen in a multicast tree and are close to the multicast source, the size of the performance cluster will be large, which adversely affects most of the group members.

Multicast tree-building algorithms employed by different overlay multicast protocols exhibit different scalability and adaptability characteristics under network dynamics during the multicast session time. A protocol may build a well-formed initial overlay multicast tree under stable network conditions, but may not be able to sustain the application performance in the presence of underlying network perturbations. Multicast protocol performance has been addressed in the context of traditionally reliable IP multicast [9]. Most of the previous overlay multicast protocols focused on the construction of overlay multicast trees. Therefore, the adaptation performance to network dynamics is either passive and limited [3], [4], [10] or not scalable [1], [2]. The adaptation is passive and limited because, although a member periodically looks for a new parent in the tree, it does not use end-to-end performance as a guide and thus may not help end-to-end application performance.

Previously, an Adaptive Overlay Multicast (AOM) that employed both end-to-end and local metrics to build the overlay multicast tree was proposed [11], [12]. Here, however, the concepts of fan-outs and foster limit in AOM are introduced and their effects on the quality of the application layer multicast tree are evaluated. Also presented here are AOM and its fault adaptation algorithm, a simulation study on AOM with different fan-outs, and a comparison of tree quality with a well-know application layer multicast tree protocol.

Related Work

The previous overlay multicast studies focused on self-organizing the group members into a delivery tree and classifying them into centralized, distributed direct-tree and distributed mesh-first approaches. ALMI [13] takes a centralized approach where a central controller builds the over-

lay and disseminates the tree information to the group members. The NARADA [2] and Gossamer [1] protocols build a mesh first and run a DVMRP-like routing protocol to build the tree. Other protocols like NICE [10] and YOID [4] build the tree directly; i.e., the tree is extended when a new joining member connects to an existing member. All these protocols use Round Trip Time (RTT) as the building metric. HostCast [14] utilizes the shortest end-to-end delay in path finding, but no effort is given to match the overlay multicast tree to the optimized IP multicast tree. None of the above protocols has investigated the loss adaptation issue.

The Host Multicast Tree Protocol (HMTP) [3] is a typical direct-tree protocol using RTT as the only metric. A new member moves as far as possible from the ROOT only if it finds a potential parent closer than the current one, and its RTT to the current parent is longer than the potential parent's RTT to its current parent. Periodically, a member randomly selects another member in its path to the ROOT (or ROOT path) and explores the branch under that member for a new parent. The periodic level-by-level exploration and probing among members accounts for most of the overhead in HMTP.

Recent tree-building approaches include closest-first-searching (CFS) [15], adjacency matrix [16] and minimum diameter spanning tree [17]. Zhang et al. use an approach similar to HMTP, except that a member tries to remember different branches in the tree building process [15]. The objective is to extend the searching range of a node position so that nearby nodes have a better chance of staying together in the tree. The algorithm itself does not consider tree maintenance or issues of adaptability. Mourad and Ahmed rely on an adjacency matrix to build the multicast tree, where matrix information is provided by the underlying P2P architecture [16]. Their application, then, must be tied into a P2P network. Moreno-Vozmediano takes a centralized approach where the multicast source node collects the probing results from every grid node and calculates the minimum spanning tree for multicast file distribution [17]. This approach is expected to be adaptive to network dynamics if continuous probing is applied; however, scalability is the main shortcoming of this centralized approach.

Chu et al. studied the overlay multicast protocol in dynamic network environments [18]. Their experiments were carried out on a mesh-first protocol, NARADA, with the results showing that it is important to adapt delay and bandwidth for conferencing applications. In this current study it was felt that it is also necessary to study the dynamic adaptation in direct-tree protocols, first because direct-tree protocols do not have an explicit multicast routing protocol—as in the case of NARADA and Gossamer—to distribute helpful information for the adaptation and, second, because one of the objectives of direct-tree protocols is scalability. A transient study can help analyze whether a protocol is scalable by adapting it to network faults efficiently and on time.

A simple, best-effort approach for improving the data delivery ratio under dynamic network conditions was recently studied in Probabilistic Resilient Multicast (PRM) [19]. The idea is that in addition to forwarding the normal data along the multicast tree, each member randomly chooses a constant number of other members and forwards the new data to each of them with a certain probability. Random forwarding sends duplicate packets to the members that are fault-free, while providing passive loss recovery at the faulty locations. PRM is not a multicast tree protocol but is a best-effort approach for improving the data delivery ratio in an overlay multicast.

An Adaptive Overlay Multicast Approach

In the following sections, fan-outs, foster limit and tree adaptation are reviewed as they relate to AOM. Additional details can be found in a study by Wu et al. [12].

AOM Tree Protocol

For scalability, the AOM tree protocol takes the direct-tree approach. The tree protocol fulfills the following tasks: tree formation, tree improvement, membership management, loop avoidance, detection and resolution. Most of the previous tree construction protocols only use Round Trip Time (RTT) to local neighbors (referred as local RTT) to connect the members. AOM, on the other hand, uses both End-to-End Delay (EED) to the ROOT and RTTs between the members to determine how to construct the tree. In Equation (1), a member's (*i*) EED to the ROOT is defined as the sum of its parent's EED and half of the RTT between the member and its parent.

$$EED_{i,ROOT} = EED_{i'sparent,ROOT} + 0.5 \cdot RTT_{i,i'sparent} \quad (1)$$

A member measures its RTT to another member by periodically sending PROBE messages. The measurements are smoothed with exponential averaging. To calculate the EED, a parent puts its current EED in the PROBEREPLY message (the ROOT's EED is 0) and a child updates its own EED, as defined in Equation (2).

$$EED_{new} = EED_{parent} + 0.5 \cdot RTT_{new} \quad (2)$$

A member joins the group by sending a JOIN message to the ROOT. If the ROOT can accommodate the new mem-

ber, it sends an ACCEPT message with the information of its current children. The new member then starts looking for the most suitable parent. In an AOM tree protocol, the more suitable parent for a member is the one that is closer to the member than the current parent (i.e., smaller RTT), is closer to the ROOT than the member (i.e., smaller EED), and through which the member's new EED is not penalized too much.

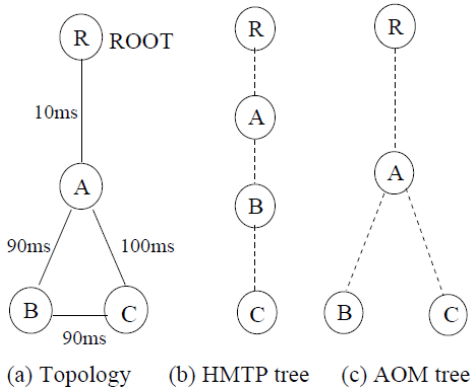


Figure 1. Example of Multicast Trees

The reason that AOM uses both EED and RTT as metrics is that EED reflects the *vertical* distance of a member to the ROOT, while RTT reflects the *local* distance between a member and its neighbor; both of them need to be considered to provide the best match between the tree and the underlying network topology. Figure 1a shows a simple network topology; Figure 1b shows a tree built by HMTP [3] that uses only RTT as the metric; Figure 1c shows a tree built by AOM, assuming the joining sequence of A, B and C.

To limit the joining overhead, a member looks for a new parent from a potential set of parents. Assuming that the ROOT is at level 0, a member is at level i if its parent is at level $i-1$. In such a case, they are said to be 1 overlay hop away from each other. The potential set of a level i member is { $member_{l,h}$ }, where $l = i$ and $h <= 4$; l is the level and h is the number of overlay hops from this member. The potential set is obtained by probing the ancestors. If a level i member finds a new parent in the potential set, its level becomes $i+1$. Parent searching continues until no new parent can be found. A nice property of such a potential set ($h <= 4$) is that when a member initially joins the group and is at level 1 or level 2, it has the opportunity to explore its position in all of the tree branches; when it moves further down the tree, however, the searching is limited to sub-branches.

Due to independent joining sequences and dynamic membership, it is necessary for the members to periodically re-evaluate their positions and continue to optimize the tree structure after joining the group. Since topologically close members are likely to stay close in the overlay by using both EED and RTT metrics, tree improvement is carried locally, i.e., a member only contacts its ancestors for improvement to reduce overhead. The ancestor set of a level i member is { $member_{l,h}$ }, where $l \in \{ i, i-1, i-2 \}$ and $h \in \{1,2,3\}$.

A single member leaving will cause the tree to become partitioned. Therefore, before a member leaves the group, it notifies its parent and children. Each child then chooses the closest ancestor (minimum RTT) or ROOT as new parent. Partitions caused by an unexpected member or link fault are detected either by the fault adaptation algorithm or by continuous loss of the PROBEREPLY messages. The simplest way to resolve the loop is to let each member attach its ROOT path information in the PROBE and PROBEREPLY messages. A member detects the loop by finding itself in the middle of its ROOT path and breaks the loop by re-joining the ROOT.

Performance Monitoring and Fault Detection

The previous direct-tree protocols, including HMTP, do not actively monitor end-to-end performance metrics. Therefore, they adapt only to local delay conditions as dictated by RTT increases. In AOM, a member monitors the performance of not only its current ROOT path, but also the paths through its ancestors (backup paths). Therefore, when a fault happens on the ROOT path, the member can select a backup path with better performance for its performance cluster. Currently, end-to-end delay and end-to-end loss rates are used as performance metrics for AOM.

A member monitors the EED on its ROOT path by periodically probing its parent. The EEDs on the backup paths are measured in the same way but less frequently because no other important information is exchanged on these paths. To prevent problems of instability, periodic measurements are smoothed with exponential averaging.

Since the ROOT path is used for data distribution, its loss rate, $l_{m,ROOT}$, can be measured by the application data. Loss rate on a backup path $l_{m,a,ROOT}$ is calculated as

$$l_{m,a,ROOT} = 1-(1-l_{a,ROOT})*(1-l_{m,a})$$

where $l_{a,ROOT}$ is the loss rate on the ancestor a 's ROOT

path and $l_{m,a}$ is the loss rate on the overlay link between the member and the ancestor, a. Since there is no application data on this link, the member asks the ancestor to periodically send a test packet.

Loss measurement of $l_{m,a}$ is a variation of the Average Loss Interval (ALI) method [20]. ALI is a better loss rate estimator than the Dynamic History Window (DHW) used by RON (resilient overlay networks) [21] and the Exponentially Weighted Moving Average (EWMA). ALI properly considers the effects of both recent and earlier loss events. Like ALI, AOM uses the weighted average loss over a few measurement intervals; but unlike in ALI, where the intervals are decided by every single loss event, the intervals in AOM are of equal lengths. This is because the two methods serve different purposes: ALI works for TCP-friendly congestion control and expects the sender to respond to every loss event; the purpose of this current study, however, was to detect the loss rate over a period of time. For details on the algorithm, please refer to the study by Wu [12].

Fault Adaptation

Without an adaptation algorithm, the only chance for protocols such as HMTP to recover from an EED fault is periodic improvement, where RTT is used to look for a closer member. Therefore, a member can bypass the EED fault if it finds a closer member not suffering the fault. Since a closer member is not necessarily an EED fault-free member, the result is random. In addition, a faulty link may affect many members' EEDs but not local RTTs, resulting in no switching efforts at all. In AOM, once the EED fault is detected, a member actively probes the ancestors for the most up-to-date EEDs and loss rates, and starts the fault adaptation algorithm, as summarized in Table 1.

It is worth pointing out the difference between tree improvement and fault adaptation. Both of them involve looking for new parents. However, the tree improvement process creates a more efficient tree, while the fault adaptation process satisfies the performance requirement.

Benefits of End-to-End Performance Monitoring

In this section, the benefits of using EED over RTT alone to adapt to network faults are presented. For better clarification, both cases are simulated in AOM. However, the case of using a local metric will apply to other protocols like HMTP. Figure 2 shows a 9-node network topology, the overlay multicast tree before the fault happens and the new overlay multicast tree when end-to-end delay is used as a fault adaptation metric in AOM. Every physical link has a delay of 10ms with the exception of link 14 which has a delay of 5ms. This is to ensure that, initially, member 4 selects member 1 as its parent in the tree. At a time of 50 seconds into the simulation, the delay of link 01 increases to 1.2 seconds, causing members 1, 3, 4, 5, 6, 7 and 8 to suffer large end-to-end delays. The total simulation time is 180 seconds.

Table 1. Fault Adaptation Algorithm

1. On detection of faults at member m:
   probe $S_a = \{ROOT, ancestors\}$ for RTTs, EEDs and loss rates.
2. Wait for reply, then update performance metrics through ancestor a in S_a as:
   $RTT_{new} = \beta \cdot RTT_{new} + (1-\beta) \cdot RTT_{old}$;
   $EED_{m,a,ROOT} = EED_{a,ROOT} + RTT_{new,a}/2$;
3. Add a to potential parent list pl if:
   $EED_{m,a,ROOT} < scale \cdot EED\_LIMIT$ &&
   $l_{m,a,ROOT} < scale \cdot LOSS\_LIMIT$
4. Find closest potential parent:
   $cur\_potential\_parent = min\_rtt(pl)$;
   if $cur\_potential\_parent == NULL$, adaptation fails, end.
   $pl = delete(pl, cur\_potential\_parent)$;
   join($cur\_potential\_parent$).
5. If not accepted by $cur\_potential\_parent$, go to step 4.

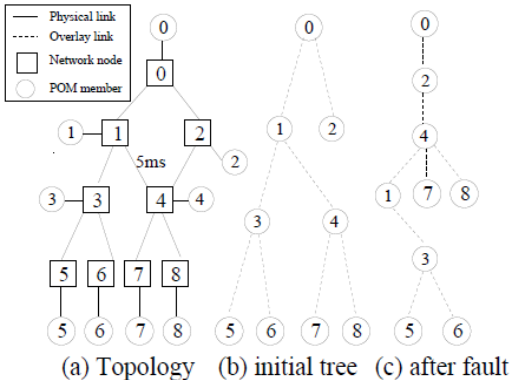


Figure 2. A 9-node Network (a), Initial Tree (b), Tree After Adaptation (c)

Figure 3 shows the simulation results of members 1, 4 and 6. It can be seen that member 4 (and, thus, its children 7 and 8) changes its path before member 1 and recovers from the fault by attaching to member 2. Member 1 could not adapt

to the fault by itself as the underlying routing algorithm happens to use the faulty link to probe member 2. However, member 4 invites member 1 after it switches its sub-cluster to member 2. From the figure, it can also be seen that member 6 adapts to the fault. However, this is not because member 6 selects a new path but rather because its grandparent, member 1, changes the ROOT path to a better position. Therefore, the sub-cluster (5 and 6) recovers from the fault without extra probing and adaptation overhead. Conversely, Figure 4 shows that if only RTT is used as the performance metric, none of these members (1, 4 or 6) could adapt to the delay fault, in spite of the existence of better paths, and the multicast tree would not change.

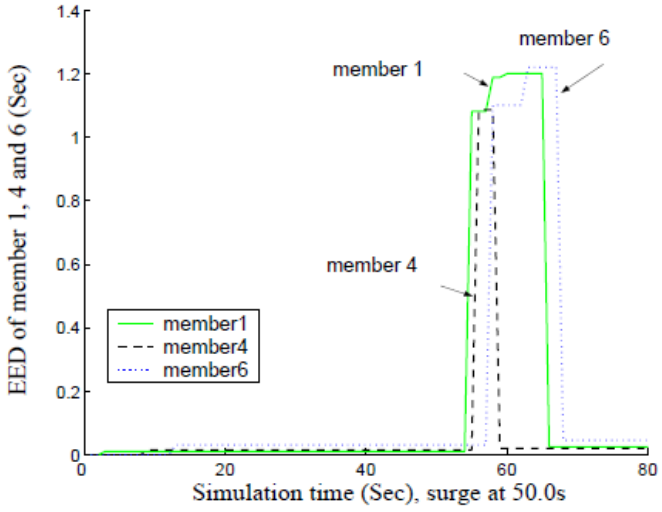


Figure 3. Adaptation of EED using End-to-End Metric

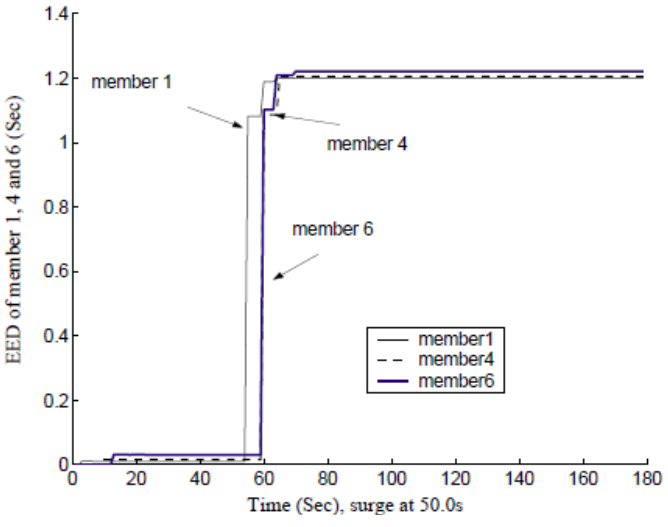


Figure 4. Adaptation of EED using Local RTT

Fan-Outs and Foster Limit

End-hosts over the Internet are heterogeneous; thus, the maximum number of unicast connections that can be set up to forward the application data depends on factors such as bandwidth capacity, traffic load and host processing power, and may vary from time to time. This connection limit is called a fan-out limit.

A member's fan-outs should be those that are best for the tree quality. If a new connection request is simply refused when the fan-out size reaches its limit, the resulting tree may be of inferior quality. To solve this problem, the connections accepted by a member are classified as fan-out connections and foster connections. Fan-out connections are used to forward application data and are restricted by the fan-out limit. Foster connections are used to construct the tree and are restricted by a foster limit. If a connection request cannot be treated as a fan-out connection, it is accepted as a foster connection for a period of time. Since the control packets used for tree construction are of small size and are sent much less frequently than data packets, a member can manage many more foster connections.

During the fostering period, several changes could occur. First, the new child may find a more suitable position in the current tree branch and move down. Second, due to periodic tree improvement, an existing child may move or become the child of a new member. Third, if none of the fan-outs or the foster child finds a better position, the tree stays unchanged. For the third case, the foster child or a fan-out will be forced to move away depending on its RTT to the parent. During the fostering period, a foster child can receive the application data from its old parent or from a randomly selected member if it is in the initial joining period. Results in the next section show how foster connections improve tree quality.

Effects of Foster Limit on Protocol Performance

This section presents the effects of foster limit on AOM performance and compares them with HMTP [3], a typical direct-tree protocol. First, tree quality is evaluated in randomly generated 1,000-node transient-stub network topologies. In this part, network conditions are static in that 1) link delays are pre-assigned and do not change during the simulations, and 2) the members do not leave the multicast group during the simulation period. The second part focuses on the adaptability of the two schemes. Faults like delay and loss-rate surges are added to randomly selected links in order to observe how the schemes respond to such events.

Performance Metrics

The following metrics are used to evaluate the tree quality:

- Group EED ratio: the group members' EED ratios are averaged. A member's EED ratio is the ratio of its overlay EED and its EED in the Shortest Path Source Tree (SPST) in an IP multicast. This metric measures the increase in EED in an overlay multicast.
- Average link stress: assuming that $LS(i)$ is the number of duplicate packets on a link i, the average link stress is defined as: $\sum_{i,LS(i)>=1} LS(i) \;/\; \sum_{i,LS(i)>=1} 1$ and reflects the load added to a link by an overlay multicast protocol.
- PDF for link stress: Link stress shows the distribution of the stress over the physical links as well as the most stressed link.
- CDF for path length: path length is defined as the number of physical links (hops) in a member's ROOT path. A longer path is not desirable because it often implies a larger EED and certainly more processing overhead.
- Tree cost: This is the total number of physical links used by the tree. Tree cost ratio is defined as the ratio of overlay multicast tree cost to the corresponding SPST tree cost.
- Control overhead: This is the total amount of control traffic used to build and maintain the tree.

Performance Evaluation

The simulation was implemented using Network Simulator-2 [22]. The 1,000-node transient-stub network topology was randomly generated by GT-ITM [23]. For simplicity, links are assigned symmetric random delays. Since both AOM and HMTP use RTT to estimate the delay between two members, symmetric delay does not favor either of them. The simulation results reflect the average of 10 runs with a C.I. of 95%, except for those that describe the transient behavior. Table 2 summarizes the values of the simulation parameters.

Table 2. Simulation Parameters

Parameters	AOM	HMTP
Improvement period	60 seconds	30 seconds
Foster limit	0, 50	0
Join time	uniform(0,1500) seconds	
Simulation time	2500 seconds	
Fan-out limit	10	
Group density	5%, 8%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%	

Figure 5 shows that for a group size ranging from 50 to 900, an AOM member has, on average, a much smaller EED than it would in HMTP. This is due to the fact that the tree algorithm in AOM considers not only the RTTs but also the EEDs to the ROOT. A member in AOM has limited tolerance to increase its EED. This avoids the long paths that could occur in HMTP, as will be shown later. With the exception of when the group size is 50, the group EED ratio of AOM is at least 60% less than that of HMTP. Another observation is that for a group size of 50 to 900, the group EED ratio in AOM remains low and stable, while in HMTP it increases by 60%. This means that the AOM tree matches the underlying network topology better and the AOM scheme is more scalable. It is also shown in Figure 5 that fostering a few members for the purpose of tree construction improves the EED ratio. At a large group size, fostering 50 children in AOM decreases the EED ratio by 15% over no fostering.

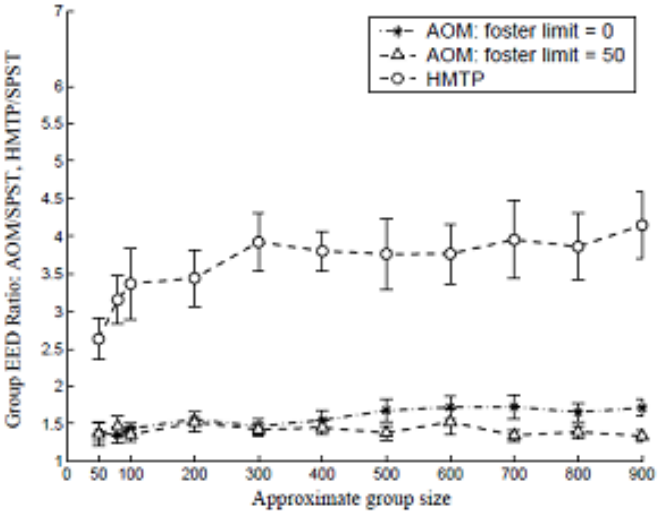


Figure 5. Tree Quality: Group EED Ratio

Unlike in an IP multicast tree where every link has a stress of 1, some links in an overlay multicast tree have duplicate packets. The link stress is affected by the group size and whether the tree matches the underlying network topology. Figure 6 shows the average link stress of HMTP and AOM. In both schemes, average link stress increases with group size. This is because the more members there are, the more likely that some links will be used repeatedly. When the group size is small, HMTP builds lower-stress trees than AOM. However, in large groups, AOM with 50 foster children outperforms both HMTP and AOM with no foster children by up to 17.36%.

Figure 7 shows the pdf of the link stresses collected from 10 simulation runs. The group density of the simulations was 85%. The largest link stress in the figure is the largest

link stress to appear in the 10 simulation runs. It can be seen that almost 99% of the links in both schemes have a stress number of less than 7; however, there are a few heavily stressed links in each scheme. In a group of about 800 members, the worst link stress to appear in 10 runs was 28 in AOM for both foster limits. HMTP had a smaller worst-link stress of 22.

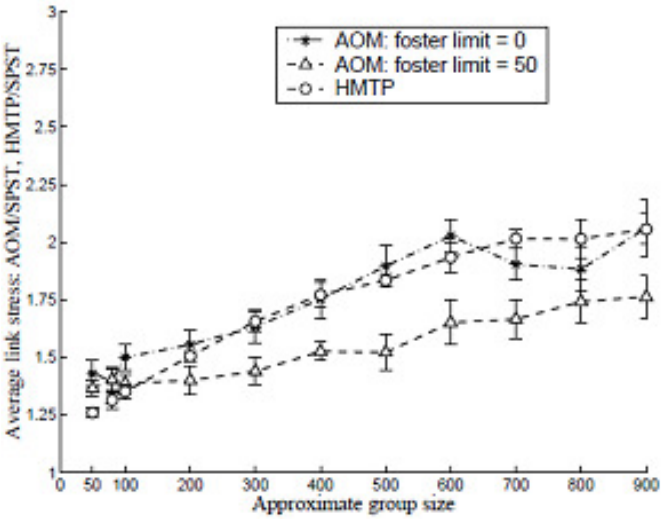


Figure 6. Tree Quality: Average Link Stress

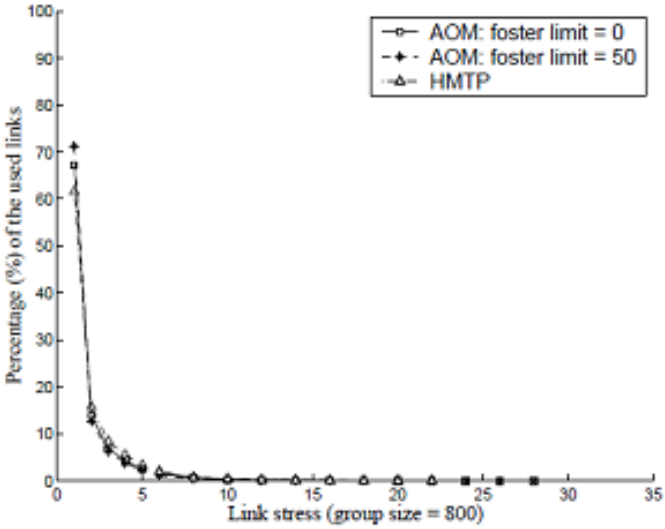


Figure 7. Tree Quality: *pdf* of Link Stress

An ideal overlay multicast tree should provide short ROOT paths and low link stress. However, it is difficult to achieve both objectives at the same time. One extreme is the multiple unicasts tree in which the paths are short but the link stresses are high. Another extreme is that the path is extremely long, but the link stress is low. Often, longer paths result in larger EEDs.

Figure 8 shows the CDF of the ROOT path length in AOM, HMTP and multiple unicasts. Path length is the number of physical links involved in a member's ROOT path. In the figure, the path lengths of all the members in a group were collected from 10 simulation runs with a group density of 85%. As can be seen, the multiple unicasts tree had the shortest path length, which was also the lower bound of the overlay multicast tree. The path length of AOM was moderate because the longest path length was 60 hops when the foster limit was 0. At least 30% of the HMTP members have a ROOT path longer than the longest path in AOM.

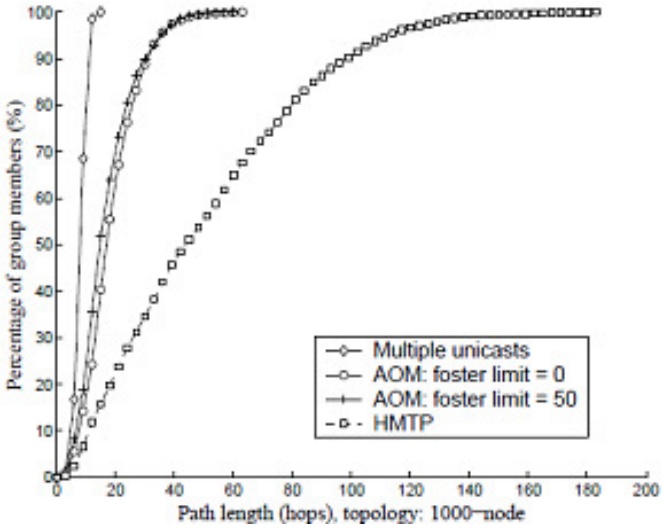


Figure 8. Tree Quality: CDF of Path Length

Tree cost reflects the total resources consumed by the overlay multicast group, such as bandwidth and processing power. Figure 9 compares the average tree cost of AOM and HMTP. For each group size, the result was the average of 10 trees and was normalized by the cost of the corresponding SPST. As can be seen in large groups, fostering children in building multicast trees in AOM saves 20%-32% more network resources than HMTP.

Figure 10 shows the change of the tree cost in a typical run of each scheme. About 800 members join the group in the first 1500 seconds. At the initial phase, the tree cost increases rapidly. After all of the members join the group, the tree cost begins to decline as the improvement algorithm continues to work. It can be seen that AOM with a foster limit of 50 had the smallest tree cost and fastest convergence.

Both AOM and HMTP need control packets to build and maintain the tree. In HMTP, control traffic is used to refresh information between parents and children, measure the RTTs and query for the information used by the tree improvement algorithm; control traffic in AOM is used to ex-

change information between a member and its ancestors. In the simulations, 40-byte control packets were used.

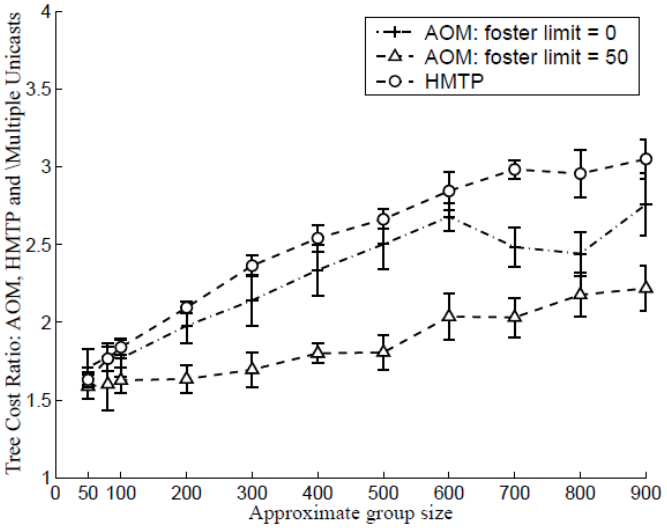


Figure 9. Tree Quality: Cost Ratio

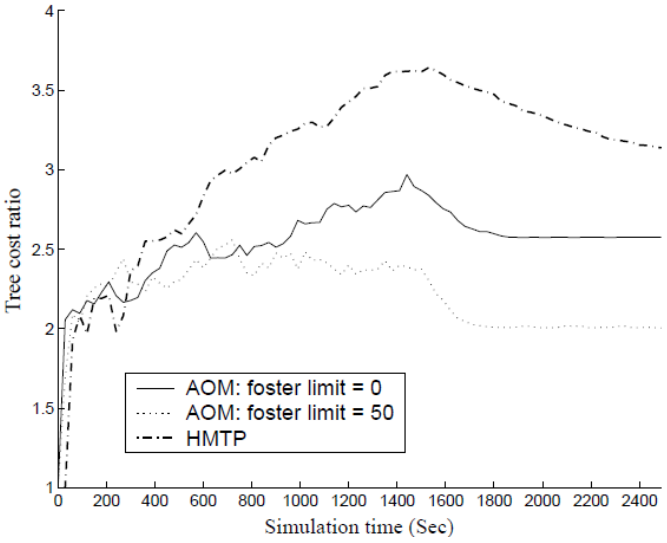


Figure 10. Convergence of Tree Cost Ratio

Figure 11 shows that the control traffic load added by the two schemes increased with the group size. At a group size of 900, the control traffic load reached 30KB/sec in HMTP but was 20% lower in AOM with a foster limit of 50 and 40% lower without fostering. Such control traffic load is not large in the sense that it is distributed across the entire network rather than on a single link. AOM with a foster limit of 50 incurs more control traffic than with no fostering.

Conclusions

End-hosts over the Internet are heterogeneous; therefore, the maximum number of unicast connections that can be set up to forward the application data (called fan-outs) depends on factors such as bandwidth, traffic load and host processing power, and may vary from time to time. This study showed that appropriate fan-out limits and the foster limit (temporary connections used for multicast tree construction rather than application data distribution) can improve overall tree quality and application performance.

Multicast applications have different performance requirements. For example, media streaming applications are sensitive to delay, loss and available bandwidth; content distribution, such as server replication and large software distribution, can be loss intolerant; a delay jitter requirement must be satisfied in voice applications. Therefore, future work with AOM should consider more end-to-end performance metrics such as available bandwidth and the jitter requirement.

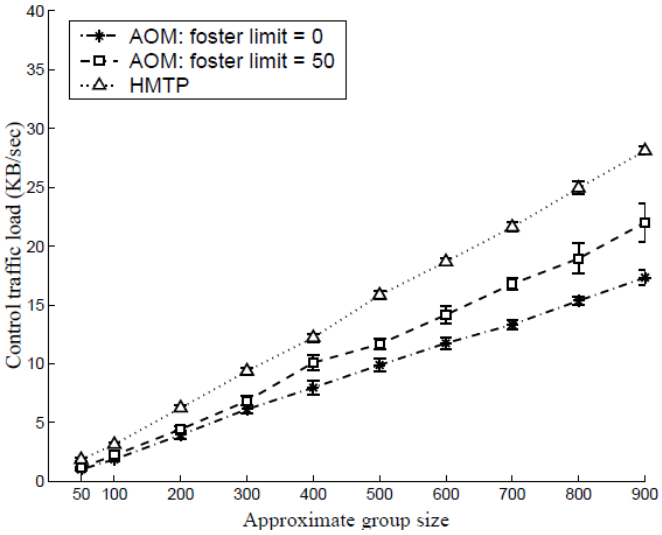


Figure 11. Overhead Traffic Load Added by the Schemes

References

[1] Chawathe, Y., McCanne, S. & Brewer, E. (2000). *An architecture for Internet content distribution as an infrastructure service*. Unpublished report available at http://www.cs.berkeley.edu/yatin/papers.

[2] Chu, Y. Rao, S. G. & Zhang, H. (2002). A case for end system multicast. *IEEE Journal on Selected Areas in Communication*, 20(8), 1455-1471.

[3] Zhang, B., Jamin, S. & Zhang, L. (2002). Host multi-

cast: a framework for delivering multicast to end users. *Proceeding of IEEE INFOCOM*.

[4] Francis, P. (2000). *Yoid: Extending the Multicast Internet Architecture*. White paper available at http://www.aciri.org/yoid/.

[5] Yajnik, M., Kurose, J. & Towsley, D. (1996). Packet loss correlation in the mbone multicast network. *Proceeding of IEEE Global Conference on Communications*.

[6] Cáceres, R., Adams, A., Bu, T., Duffield, N. G., Friedman, T., Horowitz, J. et al. (2000). The Use of End-to-End Multicast Measurements for Characterizing Internal Network Behavior. *IEEE Communications*, 38(5).

[7] Labovitz, C., Ahuja, A. Bose, A. & Jahanian, F. (2000). Delayed internet routing convergence. *Proceeding Of ACM SIGCOMM*, (pp. 175-187).

[8] NLANR active measurement project. (n.d.). Retrieved February 1, 2006, from http://watt.nlanr.net/ .

[9] Birman, K. P., Hayden, M., Ozkasap, O., Xiao, Z., Budiu, M. & Minsky, Y. (1999). Bimodal multicast. *ACM Transactions On Computer Systems*, 17(2), 41-88.

[10] Banerjee, S., Bhattacharjee, B. & Kommareddy, C. (2002). Scalable application layer multicast. *Proceedings of ACM SIGCOMM*.

[11] Wu, S. & Banerjee, S. (2004). Improving the Performance of Overlay Multicast with Dynamic Adaptation. *Proceedings of IEEE Consumer Communications and Networking Conference (CCNC)*, Las Vegas, NV.

[12] Wu, S., Banerjee, S., Hou, X. & Thompson, R. A. (2004). Active Delay and Loss Adaptation in Overlay Multicast. *Proceeding of IEEE International Conference on Communications*.

[13] Pendarakis, D., Shi, S., Verma, D. & Waldvogel, M. (2001). ALMI: an application level multicast infrastructure. *Proceedings of the 3rd USENIX Symposium on internet technologies and systems (USITS)*.

[14] Li, Z. & Mohapatra, P. (2003). HostCast: A New Overlay Multicasting Protocol. *Proceeding of IEEE International Conference on Communications*.

[15] Zhang, X., Li, X., Wang, Z. & Yan, B. (2008). A Delivery Tree Building Approach for Application Layer Multicast. *International Symposium on Computer Science and Computational Technology*, (pp. 630-634).

[16] Mourad, A. & Ahmed, M. (2008). A Scalable Approach for Application Layer Multicast in P2P Networks. *Pervasive Computing and Communications, IEEE International Conference*, (pp. 498-503).

[17] Moreno-Vozmediano, Y. R. (2009). Application Layer Multicast for Efficient Grid File Transfer. *International Journal of Computer science and Applications*, 6(1), 70-84.

[18] Chu, Y., Rao, S. G., Seshan, S. & Zhang, H. (2001). Enabling Conferencing Applications on the Internet Using an Overlay Multicast Architecture. *ACM SIGCOMM*, San Diago, CA.

[19] Banerjee, S., Lee, S., Bhattacharjee, B. & Srinivasan, A. (2003). Resilient Multicast Using Overlays. *ACM SIGMETRICS International Conference on Measurement and Modeling of Computer Systems*, San Diego, CA.

[20] Floyd, S., Handley, M., Padhye, J. & Widmer, J. (2000). Equation-based Congestion Control for Unicast Applications. *Proceeding of ACM SIGCOMM* (pp. 43-54).

[21] Andersen, D., Balakrishnan, H., Kaashoek, F. & Morris, R. (2001). Resilient Overlay Networks. *Proceeding of the 18th ACM SOSP*, Banff, Canada.

[22] The Network Simulator ns-2. (n.d.). Retrieved May 20, 2009, from http://www.isi.edu/nsnam/ns/.

[23] GT-ITM: Georgia Tech Internetwork Topology Models. Retrieved June 18, 2009, from http://www.cc.gatech.edu/fac/Ellen.Zegura/graphs.html.

Biographies

XIAOBING HOU is currently an Assistant Professor at the Computer Electronics and Graphics Technology Department at Central Connecticut State University. He received the Ph.D. degree in Information Science from the University of Pittsburgh in 2006. Dr. Hou's teaching and research interests are in the area of computer networking and information security. He is a member of IEEE and ACM. Dr. Hou may be reached at xhou@ccsu.edu

SHUJU WU is currently an Associate Professor at the Computer Electronics and Graphics Technology Department at Central Connecticut State University. She received her Ph.D. degree in Information Science from the University of Pittsburgh in 2004. Dr. Wu's teaching and research interests include computer communications and networks, multimedia systems, performance modeling and evaluation, and network applications. Dr. Wu may be reached at swu@ccsu.edu

DIGITAL BREAKTHROUGH DETECTION USING LASER-INDUCED, THERMAL DIFFUSION SHOCK WAVES

Jun Kondo, University of Hartford; Saeid Moslehpour, University of Hartford; Hisham Alnajjar, University of Hartford

Abstract

The efficiency of a jet engine is improved by increasing the temperature in the engine combustion components. Combustion chamber temperatures have increased up to 1600°C over the past decade [1]. Therefore, jet engine combustion components must deal with these increased temperatures. Free-air-flow cooling holes are critical for cooling the components, but the process of drilling cooling holes presents numerous problems. The main problem to be addressed is "back wall strike". This study looked at innovative approaches to designing controllers for the laser percussion drilling process to determine the exact moment of breakthrough that could eliminate back wall strike, which damages the adjacent surface of jet-engine turbine components. The PCB 106B pressure sensor was used to measure thermal diffusion shock waves, and National Instruments LabVIEW computer program was used to establish control algorithms. The controllers process the sensor output digitally to determine the exact moment of breakthrough, thereby eliminating back wall strike. There were two methods for processing the sensor output digitally: software and hardware. In the software method, LabVIEW was used to extract pulse signal components from the sensor output and the laser power output. In the hardware method, operational amplifiers were used to extract pulse signal components from the sensor output and the laser power output. The processed sensor output showed distinctive patterns, which indicated the relationship between the laser pulse and the shock pulse at the moments of breakthrough. Therefore, the system successfully detected the breakthrough using the digital approach.

Introduction

The laser percussion drilling process at the Connecticut Center for Advanced Technology (CCAT) is shown in Figure 1. The laser beam was generated by the neodymium-doped yttrium aluminum garnet (Nd: YAG) laser of the Convergent Prima P-50 laser drilling machine at CCAT. The laser beam passed through the center of the copper nozzle and impinged upon the surface of a Waspalloy steel plate sample. The angle between the laser and sample was 20 degrees, which is the standard for cooling-hole drilling for jet engine turbine blades. After a few percussion drilling operations, the laser beam started penetrating the sample and making a small diameter hole on the sample surface; this process is known as partial breakthrough. At the next laser shot, the laser beam completely penetrated the sample; this process is known as full breakthrough. But subsequent laser shots continuously drilled the adjacent sample surface after full breakthrough in the laser percussion drilling process of actual jet engine turbine blades. This unavoidable process is known as back wall strike. In order to diminish the effect of back wall strike, Loctite Hysol 7901 polyamide hot melt might be injected in cavities of jet engine turbine blades. But the adjacent sample surface might receive serious surface damage despite the existence of the hot melt. In order to solve this problem, the exact moment of full breakthrough must be detected by the sensor, and the controller must turn off the laser immediately at the exact moment of full breakthrough.

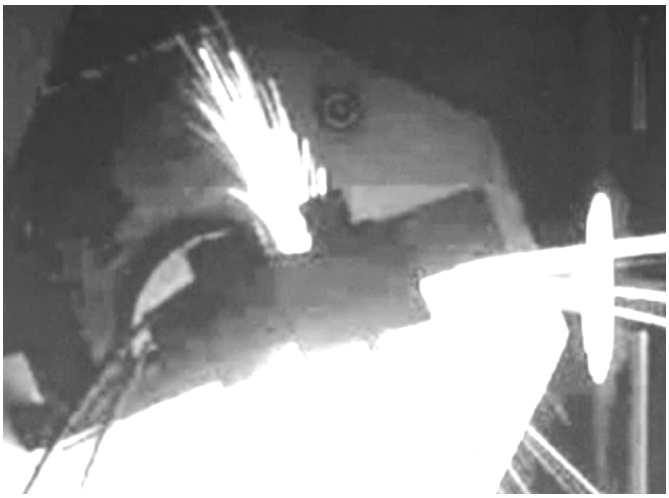


Figure 1. Laser Percussion Drilling Process at Connecticut Center for Advanced Technology (CCAT)

Many approaches have been developed to minimize the effect of back wall strike. Full breakthrough can be detected by frequency changes of the drilling sound signatures using Fast Fourier Transform (FFT). It can also be detected by spectrum changes of the percussion drilling arc. Another possibility is detection by a video camera, which would be mounted to view the area being drilled through a path coaxial with the drilling laser beam [2]. In this project, the PCB-106B pressure sensor was used to measure Laser-Induced Thermal Diffusion Shock Waves to examine the thermal

contact between the laser beam and the turbine blade to detect the exact moment of full breakthrough. The output of the PCB-106B pressure sensor and the output of the laser power sensor were digitized using the software method or the hardware method to produce the shock pulse and the laser pulse. Finally, the shock pulse was subtracted from the laser pulse to detect the exact moment of full breakthrough when the laser beam completely penetrated the sample.

Related Research

The effects of Laser Induced Thermal Diffusion Shock Waves have been investigated and the fundamental equations were established by Danworaphong et al. [3] in the book "Laser Induced Thermal Diffusion Shock Waves." When a neodymium-doped yttrium aluminum garnet (Nd: YAG) laser induces a thermal diffusion shock wave, the thermodynamic properties−speed U, density r, and pressure P−are dramatically different before the shock front and after the shock front. The figure of the shock front is shown in Figure 2.

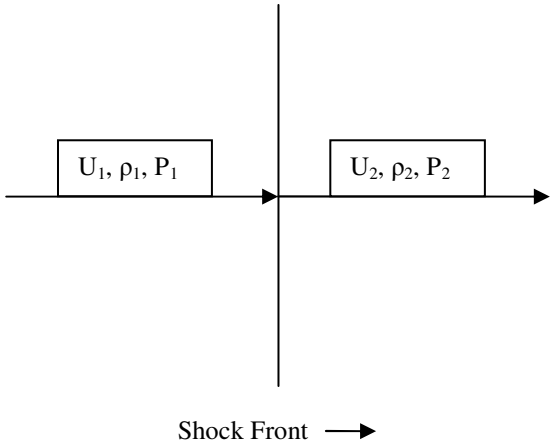


Figure 2. "Shock Front"

Thermal diffusion shock waves have several properties identical to fluid shock waves generated by supersonic flight [4]. The difference between thermal diffusion shock waves and fluid shock waves is as follows [4]:

1. Thermal diffusion shock waves depend on the existence of externally imposed temperature gradients, while fluid shock waves have no such requirement.
2. Thermal diffusion shock waves always appear as a pair of identical shock fronts that propagate in opposite directions.
3. The dissipating force is viscous damping and mass diffusion in thermal diffusion shock waves. Therefore, the speed of thermal diffusion shock waves will eventually be equal to zero even in the absence of mass diffusion.

The thermal diffusion shock waves and the mass diffusion shock waves are governed by the following equation [4]:

$$\frac{\partial_{C(z,t)}}{\partial_{\tau}} = \alpha \frac{\partial}{\partial_z}\{c(z,t)[1-c(z,t)]\cos z\} + \frac{\partial^2 c(z,t)}{\partial z^2} \quad (1)$$

The significance of this equation is stated as follows [4]:

1. The first term corresponds to thermal diffusion shock waves, while the second term corresponds to mass diffusion shock waves.
2. The sinusoidal function governs the first term that represents thermal diffusion shock waves.
3. α is the thermal diffusion factor that governs the dominance of thermal diffusion shock waves over mass diffusion shock waves and is expressed as

$$\alpha = \frac{D' T_0}{D} \quad (2)$$

where

D = Mass Diffusion Constant
D' = Thermal Diffusion Constant
To = Temperature.

Partial Breakthrough and Full Breakthrough

In the percussion drilling process, the laser beam was generated by the neodymium-doped yttrium aluminum garnet (Nd: YAG) laser. It passed through the center of the copper nozzle and impinged upon the surface of a Waspalloy steel plate sample. It penetrated the sample after repeated drilling and made a small diameter hole. This condition is called partial breakthrough. At the following laser shot, the laser beam completely penetrated the sample and made a large-diameter hole. This condition is called full breakthrough. These conditions are shown in Figure 3. The diameters of these holes can be estimated using the diameter of calibration dots.

Methodology

The laser percussion drilling process setup at CCAT is shown in Figure 4. The laser beam was generated by the neodymium-doped yttrium aluminum garnet (Nd: YAG) laser of the Convergent Prima P-50 laser drilling machine at CCAT. The laser beam passed through the center of the copper nozzle and impinged upon the surface of a Waspalloy steel plate sample. The thermal diffusion shock waves

were measured by the PCB-106B pressure sensor that was placed under the sample. Also, the penetrating laser power was measured by the breakthrough detector that was placed above the sample in order to confirm the moment of breakthrough that was detected by the PCB-106B pressure sensor.

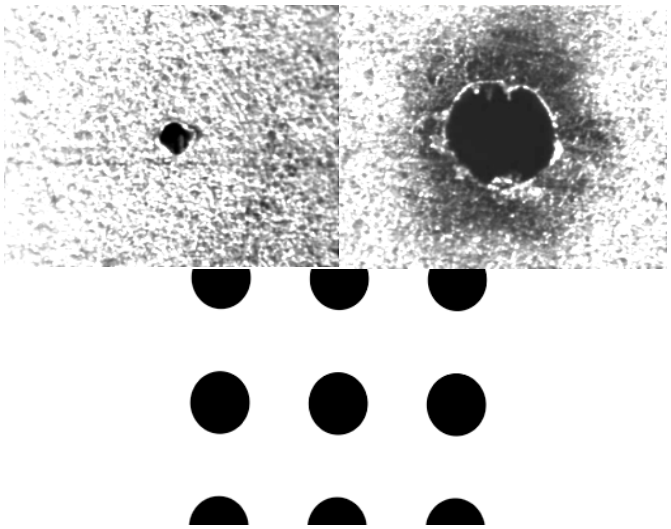


Figure 3. Partial Breakthrough, Full Breakthrough and Calibration Dots (0.25mmØ)

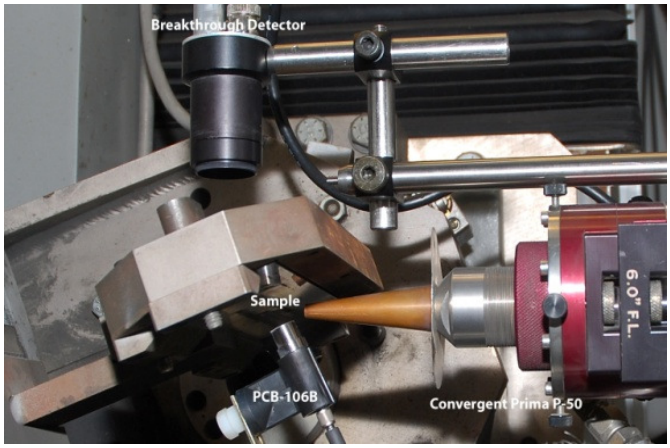


Figure 4. Laser Percussion Drilling Process Setup at Connecticut Center for Advanced Technology (CCAT) Breakthrough Detector (top) PCB 106B Pressure Sensor (bottom)

After full breakthrough, subsequent laser shots continuously drilled the adjacent sample surface in the actual laser percussion drilling process, which became the major problem to be solved. In order to eliminate the effect of back wall strike, the exact moment of full breakthrough had to be detected by processing the output of the PCB-106B pressure sensor, and the controller had to turn off the laser immediately after the exact moment of full breakthrough in order to prevent the excessive laser drilling process that damages the adjacent sample surface.

Apparatus

The National Instruments PXI-4462 Dynamic Signal Acquisition Device and the LabVIEW breakthrough detection program were used for the digital approach; the National Instruments PXI-4462 Dynamic Signal Acquisition Device is shown in Figure 5. The vertical line of the PCB106B pressure sensor output was extracted and digitized. This digital signal is called shock pulse. The laser power was also digitized. This digital signal is called laser pulse. The shock pulse was subtracted from the laser pulse in order to detect the moment of breakthrough. This process is shown in Figure 6. There are two methods to process the output of the PCB-106B pressure sensor in the digital approach: software and hardware.

Figure 5. National Instruments PXI-4462 Dynamic Signal Acquisition Device (the first module from the right) and PXIe-1062Q PXI Express Chassis

Software Method

The LabVIEW breakthrough detection program for the digital approach is shown in Figure 7. The top row of three Express VIs represents the pressure sensor block diagram that produces the shock pulse. The bottom row of three Express VIs represents the laser power block diagram that produces the laser pulse. In order to detect breakthrough, the shock pulse was subtracted from the laser pulse using the subtraction block, which is in the upper middle of the program. Also, the program recorded the following three signals in the TDMS format and saved the data on the hard drive:

1. Shock Pulse
2. Laser Pulse
3. Breakthrough Detection Signal

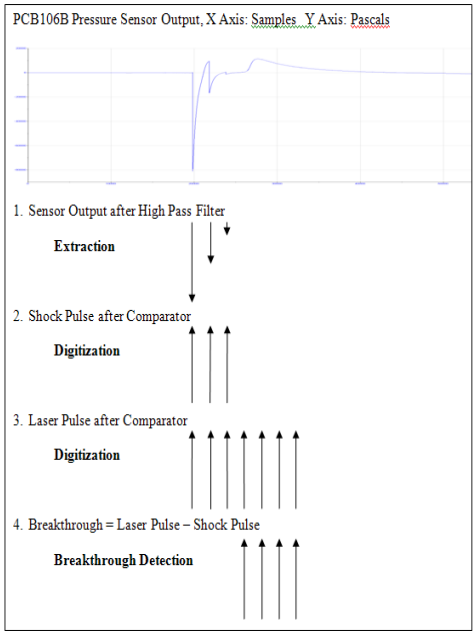


Figure 6. Breakthrough Detection Process using the Digital Approach

Hardware Method

Breakthrough Detection Circuit

The Cadence Allegro Design Entry CIS breakthrough detection schematic is shown in Figure 8. The top three rows of the operational amplifiers are the pressure sensor circuits that produce the shock pulse. The lower three rows of the operational amplifiers are the laser power circuits that produce the laser pulse. In order to detect breakthrough, the shock pulse was subtracted from the laser pulse using the subtraction circuit that is the far right operational amplifier circuit. The dual differential comparator for the pressure sensor circuits consists of six operational amplifiers, which are the left three operational amplifiers in the second and third rows. The dual differential comparator for the laser power circuits also consists of six operational amplifiers, which are the left three operational amplifiers in the fifth and sixth rows.

The schematic circuit of Figure 8 includes the following:

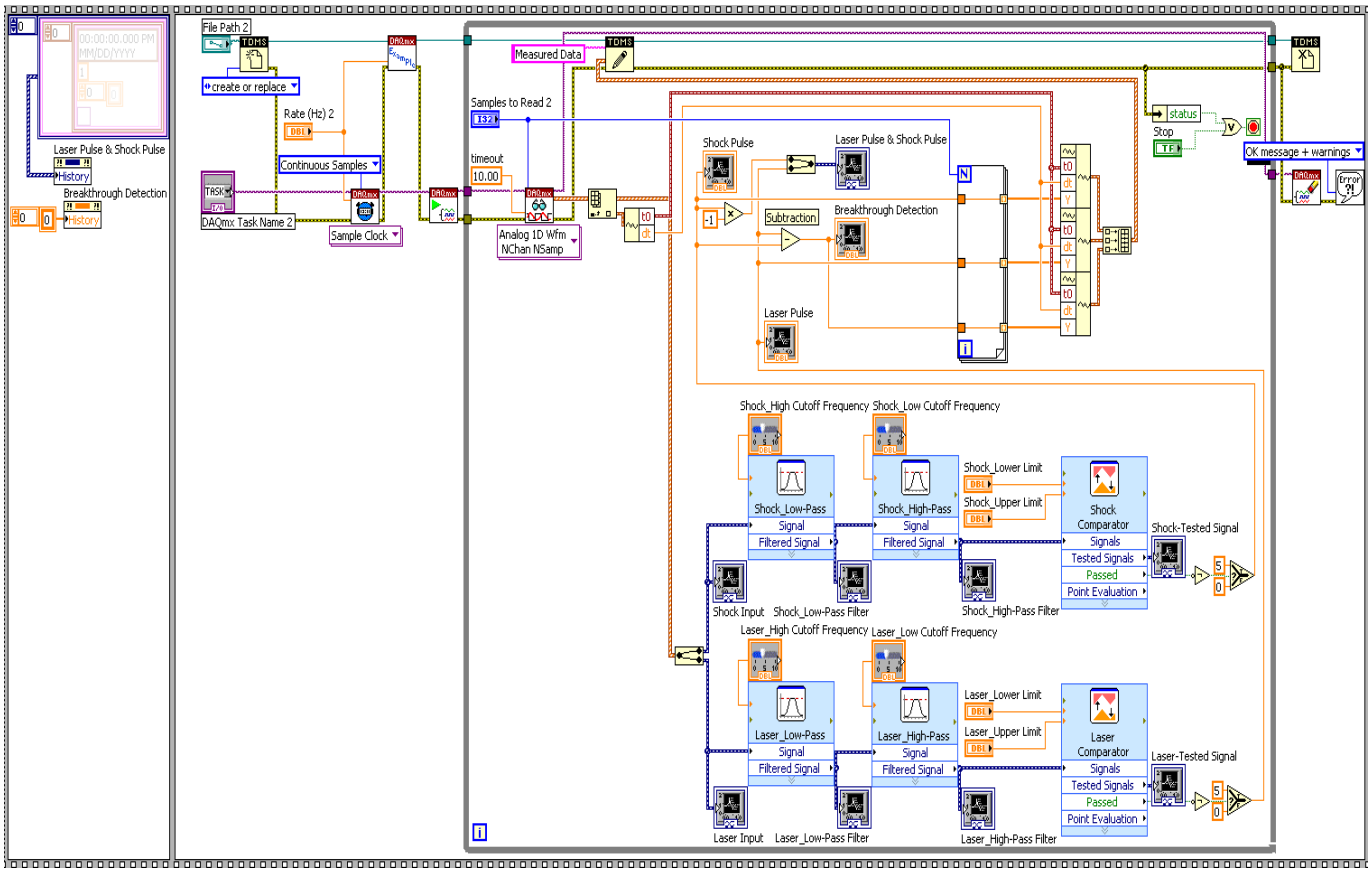


Figure 7. LabVIEW Breakthrough Detection Program for the Digital Approach

- Pressure sensor circuit (1)
- Unity-gain buffer (2)
- Low-pass filter (3)
- High-Pass Filter)4)
- Comparator (5)
- Summation (6)
- Inverter Laser Power Circuit (7)
- Unity-Gain Buffer (8)
- Low-Pass Filter (9)
- High-Pass Filter (10)
- Comparator (11)
- Summation (12)
- Inverter Subtraction Circuit (13)

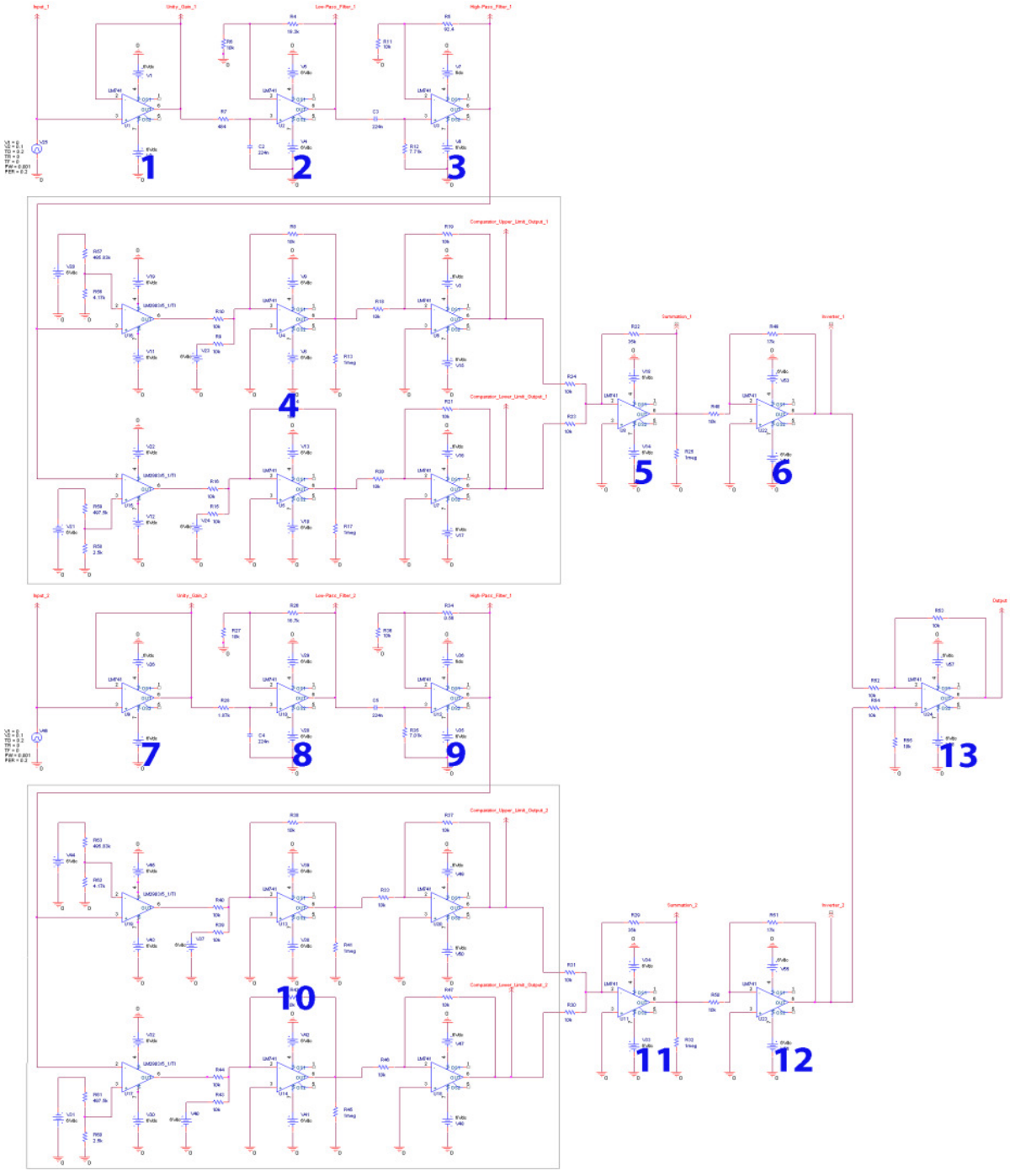


Figure 8. Cadence Allegro Design Entry CIS Breakthrough Detection Schematic

The actual circuit is shown in Figure 9. There are thirteen integral circuits. Eleven of them are operational amplifiers (NTE941M) and two of them are comparators (Texas Instruments LM2903P).

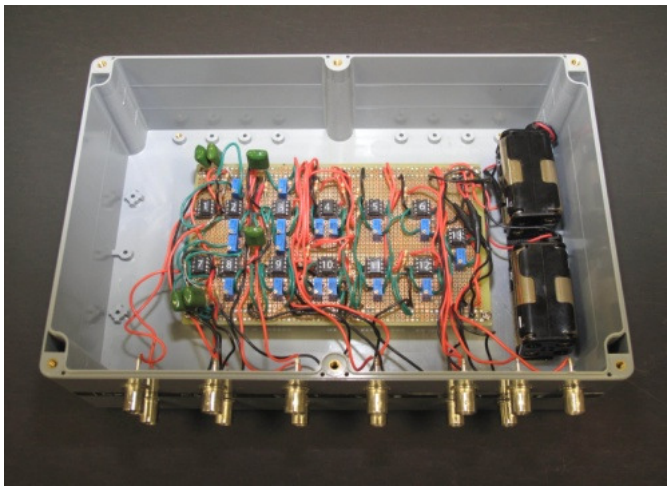


Figure 9. Breakthrough Detection Circuit

Testing Schematic

In order to test the design for the digital approach hardware method, the CIS program was exported to National Instruments Multisim, where it was simulated using the recorded TDMS file. The exported Multisim program is shown in Figure 10, and the simulation results are shown in Figure 11.

Testing Hardware

The breakthrough detection circuit was tested using the recorded data, and the data were recorded to the TDMS file using the National Instruments PXI-4462 Dynamic signal acquisition device and the LabVIEW breakthrough detection program. A sampling rate of 10kHz was used to record the data. The PCB106B pressure sensor signal and the laser pulse signal were extracted from the original TDMS file to produce the new TDMS file. This new TDMS file was played back by the LabVIEW TDMS file playback program, shown in Figure 12, to test the breakthrough detection circuit. This program has the following features:

1. It can play back the TDMS files that are recorded using any sampling rates.
2. It can play back two channels in the TDMS file simultaneously for comparison.
3. It can output the signal to any sound cards to produce the analog output waveform.

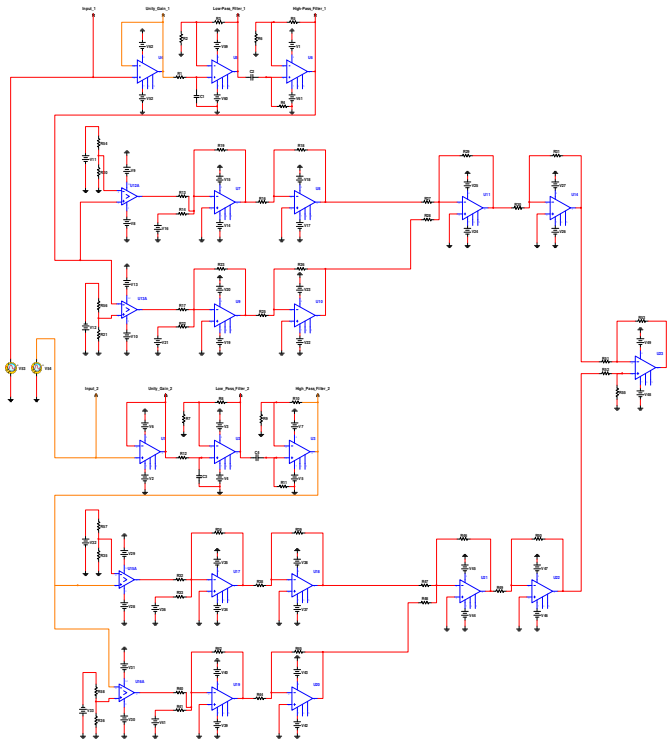


Figure 10. National Instruments Multisim Breakthrough Detection Schematic

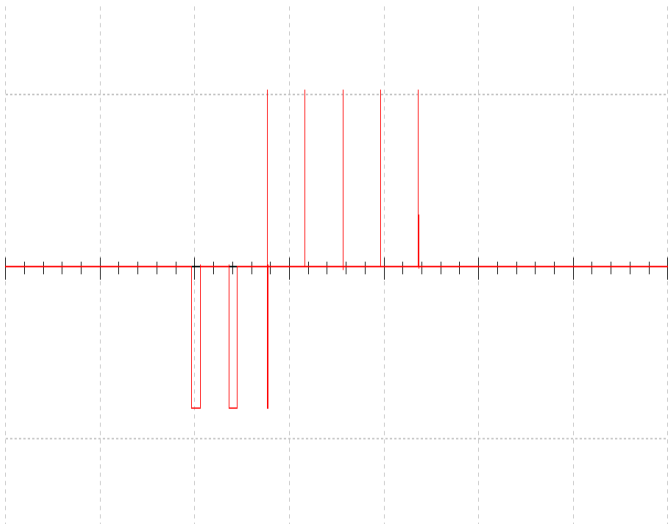


Figure 11. National Instruments Multisim Simulation Results. X Axis: Time in 500 ms/division Y Axis: Sensor Output in 5 volts/division

The Creative Sound Blaster X-Fi Titanium sound card was used to produce the analog input for the breakthrough detection circuit. The sound card can produce a fairly accurate analog signal compared to the original digital signal because of the 16-bit digital-to-analog conversion and the PCI Express bus connection.

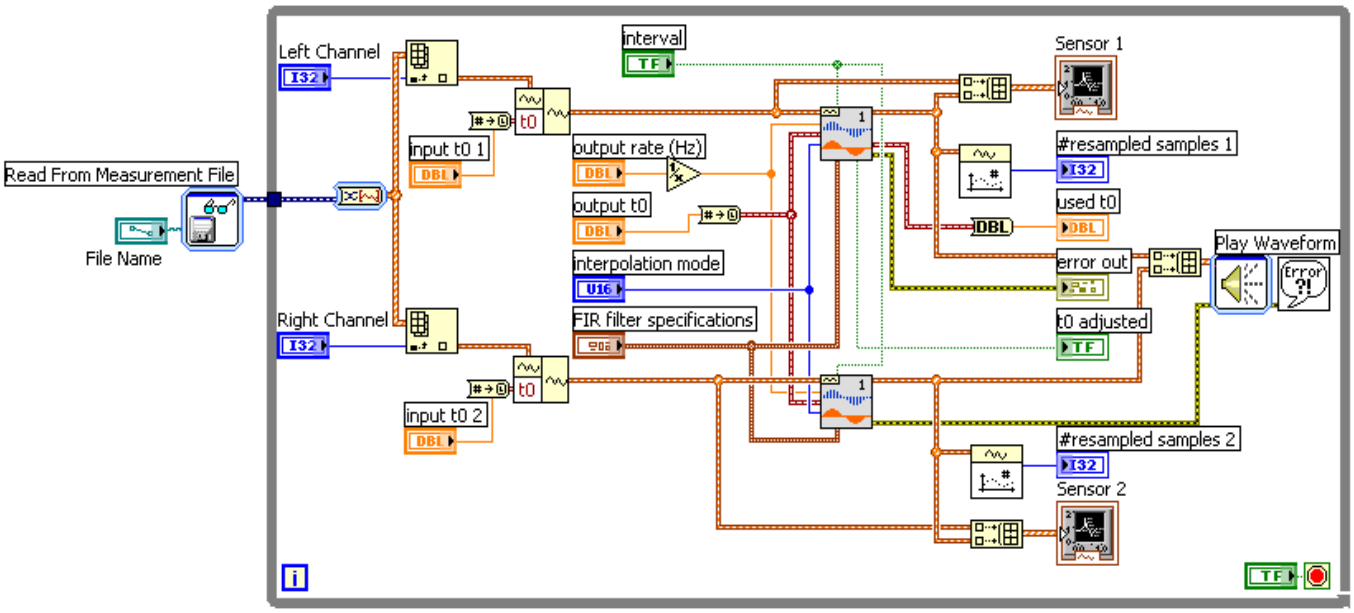


Figure 12. LabVIEW TDMS File Playback Program

Results

Software Method: Laser Pulse and Shock Pulse

Figure 13 shows the laser pulse and the shock pulse from the results. The top red line indicates the laser pulse and the bottom blue line indicates the shock pulse. The third laser shot produced partial breakthrough, while the fourth laser shot produced full breakthrough. But the third shot did not indicate partial breakthrough using this method, but rather that the laser did not penetrate the sample (refer to Figure 15 of the hardware method). The hardware result has much higher resolution and clearly shows the moment of partial breakthrough as the time delay between the laser pulse and shock pulse at the third shot.

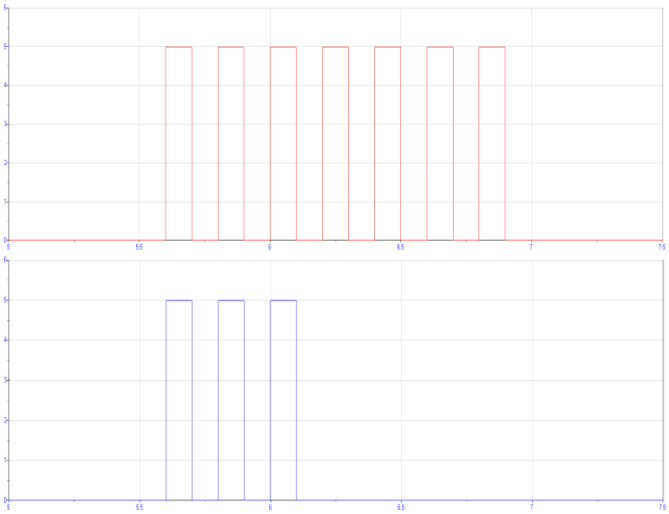


Figure 13. Laser Pulse (top) and Shock Pulse (bottom) of the Software Method. X Axis: Time in Second Y Axis: Sensor Outputs in Voltage

Software Method: Breakthrough Detection

In order to determine the moment of breakthrough, the shock pulse was subtracted from the laser pulse. Figure 14 shows the results of this subtraction or breakthrough detection. The descriptions of seven laser shots are as follows:

1. The first shot did not appear. The shock pulse was subtracted from the laser pulse, thus that result was zero.
2. The second shot also did not appear. The shock pulse was subtracted from the laser pulse, again resulting in zero.
3. The third shot also did not appear. The shock pulse was subtracted from the laser pulse yielding, again, a result of zero. Therefore, the third shot did not indicate partial breakthrough (refer to Figure 16 to compare the results of the software method to the hardware method).

4. The fourth shot was positive and indicated full breakthrough. The negative component, which was the shock pulse, completely disappeared and the positive component, which was the laser pulse, kept appearing.

5. The fifth shot was also positive. The laser beam cleaned up the existing hole.

6. The sixth shot was positive. The laser beam further cleaned up the existing hole.

7. The seventh shot was positive. The laser beam further cleaned up the existing hole.

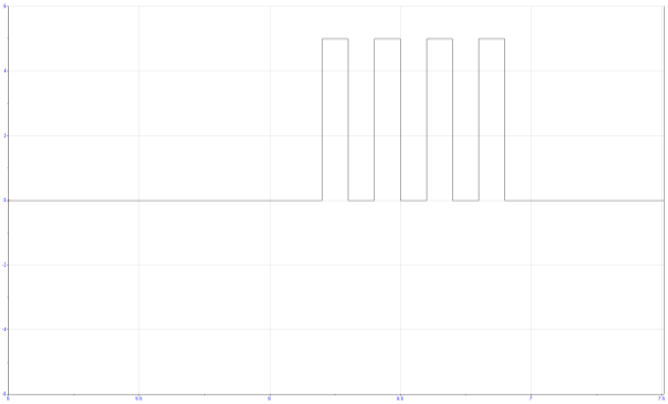


Figure 14. Breakthrough Detection of the Software Method (only fourth, fifth and sixth shots appeared)
X Axis: Time in Second Y Axis: Output in Volts

Hardware Method: Laser Pulse and Shock Pulse

Figure 15 shows the laser pulse and the shock pulse from the results. The top red line indicates the laser pulse and the bottom blue line indicates the shock pulse. The third laser shot produced partial breakthrough, where the fourth laser shot produced full breakthrough.

Hardware Method: Breakthrough Detection

In order to determine the moment of breakthrough, the shock pulse was subtracted from the laser pulse. Figure 16 shows the results of this subtraction or breakthrough detection. The descriptions of seven laser shots are as follows:

1. The first shot was negative. The negative component, which was the shock pulse, fully appeared. It indicated that drilling was in progress.

2. The second shot was also negative. The negative component, which was the shock pulse, appeared again. It indicated that drilling was still in progress.

3. The third shot was both positive and negative and produced partial breakthrough. The negative component, which was the shock pulse, partially appeared and the positive component, which was the result of the subtraction, started appearing.

4. The fourth shot was positive and produced full breakthrough. The negative component, which was the shock pulse, disappeared and the positive component kept appearing.

5. The fifth shot was also positive. The laser beam cleaned up the existing hole.

6. The sixth shot was positive. The laser beam further cleaned up the existing hole.

7. The seventh shot was positive. The laser beam further cleaned up the existing hole.

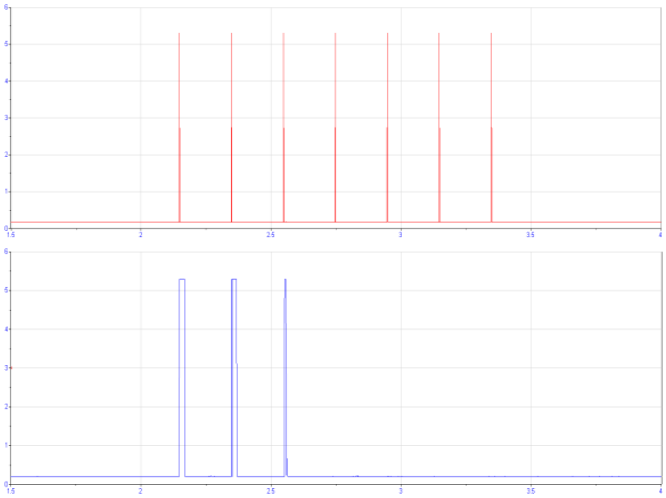


Figure 15. Laser Pulse (top) and Shock Pulse (bottom) of the Hardware Method. X Axis: Time in Second Y Axis: Sensor Outputs in Voltage

Discussion

Pressure Sensor versus Microphone

The PCBD20 ICP array microphone had been used from 2006, but was damaged by high pressure caused by the percussion drilling process in the summer of 2007. Therefore, the PCB106B series pressure sensors were recommended by PCB engineers. They decisively said that pressure caused by

the percussion drilling process was beyond the microphone's measurement range. A system based on a microphone is inappropriate for the percussion drilling process because the maximum pressure reaches 81.099kPa at 1 inch from the sample. This pressure is approximately 80 percent of the theoretical pressure limit of 101.325kPa at 1 atmosphere environmental pressure [6]. Even if the distance is increased twice to decrease the pressure to 20.275kPa, it is still over the allowable maximum pressure, 15.9kPa, of the PCB377A12 microphone that has a sensitivity of 0.25mV/Pa [5]. In addition, the PCB377A12 does not provide high sensitivity for the laser-induced thermal diffusion shock waves as does the PCB106B pressure sensor. The PCB377A12 microphone is one of the lowest sensitivity microphones made by PCB and is used in a high-pressure environment. Therefore, the pressure sensor must be used in the laser percussion drilling process at CCAT to provide both the high-pressure resistance and the high sensitivity for the laser-induced thermal diffusion shock waves to establish a consistently reliable control system that works under any conditions.

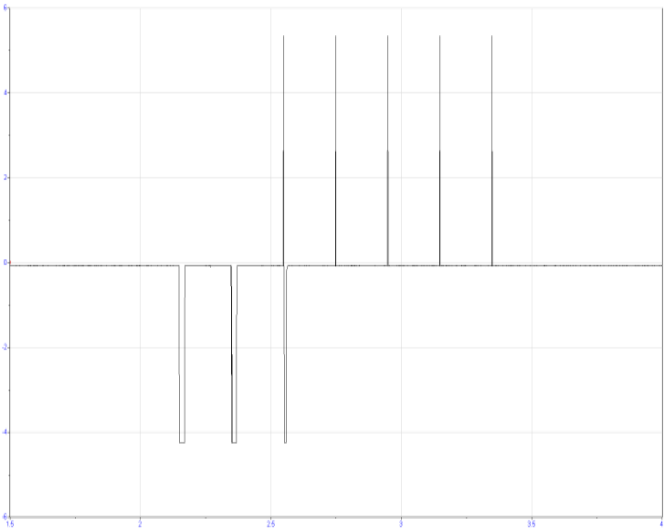


Figure 16. Breakthrough Detection of the Hardware Method X Axis: Time in Second Y Axis: Output in Voltage

Cleanup Shots

After full breakthrough, the re-solidified material might be left in the hole. A photograph of re-solidified material is shown in Figure 17. The size of it can be estimated using the diameter, 0.25mm, of the calibration dots. In order to take out the re-solidified material from the hole, cleanup shots are required after full breakthrough. But cleanup shots also continuously drill the adjacent sample surface after full breakthrough. The dilemma, then, is whether or not to continue the laser shots. Therefore, the minimum amount of laser power should be used for cleanup shots after full breakthrough.

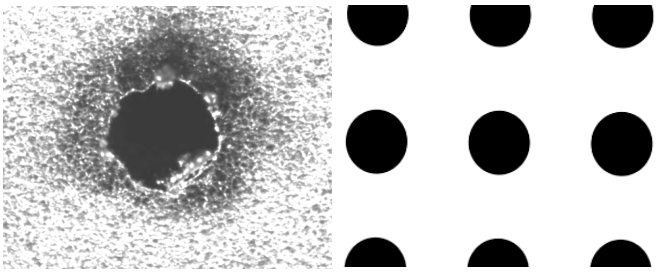


Figure 17. Re-solidified Material and Calibration Dots (0.25mm Ø)

Next Research Phases

The first fundamental experiments were accomplished in a limited time period to prove that this method is feasible. In the actual percussion laser drilling process, the turbine blade would always be rotating and all parameters continuously changing. But the exact moment of breakthrough has to be determined regardless of these unsteady conditions. Therefore, tests will be conducted under the following conditions.

1. The angle between the laser and the sample:
   Because the 20-degree laser shot is the standard for cooling hole drilling for jet engine turbine blades, this laser angle shot was used in this project. Varieties of angles will be tested to establish a consistently reliable control system.

2. The thickness of the sample:
   The thickness of the sample is significant because the 20-degree shot is the standard, and the laser beam has a relatively long distance to penetrate at this angle. But the Waspalloy samples tested here only had a thickness of 0.05 inches. Therefore, thicker samples will be tested in the future.

3. The coating of the sample:
   It is known that the thermal coating on the sample surface dramatically increases the sound signature. But coated samples have not been tested in this project. Therefore, thermal coated samples will be tested in the future.

Summary

In the digital approach, the PCB106B pressure sensor output showed distinctive patterns, which indicated the rela-

tionship between the laser pulse and the shock pulse, as shown in Figure 16. Therefore, the system successfully detected the moments of breakthrough using the digital approach. Also, these results showed that the digital approach had unique advantages and disadvantages. For example, it resulted in distinctive patterns that indicated the relationship between the laser pulse and the shock pulse. But the circuit required precise calibrations for inductance, capacitance and resistance values. Because all drilling conditions are constantly changing during actual fabrication of jet engine turbine blades, it must be tested under many different conditions to establish a consistently reliable control system that works under any conditions.

References

[1] Verhoeven, K. (2004). *Modeling Laser Percussion Drilling*. Eindhoven University, Netherlands. http://alexandria.tue.nl/extra2/200412856.pd

[2] Vanderwert, T. (2006). Breakthrough Detection. *Industrial Laser Solution*, 12. Retrieved from http://www.industrial-lasers.com/articles/2006/06/breakthrough-detection.html

[3] Danworaphong, S., Diebold, G. & Craig, W. (2008). *Laser Induced Thermal Diffusion Shock Waves*. VDM Verlag Dr. Műller. Saarbrűcken, Germany.

[4] Danworaphong, S., Craig, W., Gusev, V. & Diebold, G. (2008). *Thermal Diffusion Shock Waves.* Brown University , Providence, RI. Retrieved from http://www.math.mcmaster.ca/craig/SoretIIBNew.pdf

[5] "PCB Piezotronics," PCB, Depew, NY. http://www.pcb.com/products/

[6] "Sound Pressure," Wikipedia http://en.wikipedia.org/wiki/Sound_pressure

Biographies

JUN KONDO received his Master of Engineering in Mechanical Engineering from University of Hartford in 1998, his Master of Engineering in Electrical Engineering from University of Hartford in 2010. Presently, he is pursuing his Ph.D. degree in Electrical Engineering at the University of Connecticut. He may be reached at kondo@hartford.edu

SAEID MOSLEHPOUR is an Associate Professor in the Electrical and Computer Engineering Department in the College of Engineering, Technology, and Architecture at the University of Hartford. He holds PhD (1993) from Iowa State University and Bachelor of Science, Master of Science (1990), and Education Specialist (1992) degrees from University of Central Missouri. His research interests include logic design, CPLDs, FPGAs, electronic system testing and distance learning. He may be reached at moslehpou@hartford.edu

HISHAM ALNAJJAR is Professor of Electrical and Computer Engineering at the University of Hartford, Connecticut (USA), where he is also the Associate Dean of the College of Engineering, Technology, and Architecture (CETA). Before that he served for nine years as the Chair of the Electrical & Computer Engineering Department at the University of Hartford. Ph.D. from Vanderbilt University, M.S. from Ohio University. His research interests include sensor array processing, digital signal processing, power systems in addition to engineering education. He may be reached at alnajjar@hartford.edu

ENABLING LARGE-SCALE PEER-TO-PEER STORED VIDEO STREAMING SERVICE WITH QOS SUPPORT

Masaru Okuda, Murray State University

Abstract

The aim of this study was to enable large-scale, high-volume, stored-video streaming service over the peer-to-peer (P2P) network with Quality of Service (QoS) support. The primary focus of this paper is to address the following technical challenges associated with the distribution of stored streaming video through P2P networks: 1) allow peers with limited transmit bandwidth capacity to become contributing sources; 2) enable discovery of time-changing and time-bounded video frame availability at participating peers; and, 3) minimize the impact of distribution source losses during video playback.

To meet the above requirements, a new video distribution model was proposed which is a hybrid between client-server and P2P. In this model, the total length of a video is divided into a sequence of small segments. The peers execute a noble scheduling algorithm to determine the order, the timing and the rate of segment retrievals from other peers. The model employs an advertisement scheme that allows the discovery of video segment availability at other peers by incorporating parameters of the scheduling algorithm. An accompanying QoS scheme can reduce the number of video playback interruptions, while one or more video sources depart from the distribution network prematurely. The simulation study confirmed that the QoS scheme over the proposed distribution network was effective against excessive network delays, including multiple instances of distribution source losses.

Introduction

With the advent of recent technological advances, multimedia streaming service over the Internet is gaining increasing importance. Video streaming applications, such as high-definition video streaming and on-line DVD and Blu-ray rentals, have a potential to enrich our lives and create new business opportunities. As high performance end-user systems are becoming widely available and the number of subscribers to high-speed Internet access services is rapidly increasing, a new computing paradigm known as a Peer-to-Peer (P2P) network has emerged. P2P enables direct exchange of contents among a group of end users without the need for a centralized management structure. P2P offers a framework in which a large-scale, distributed and self-organizing content distribution network (CDN) can be constructed.

Although P2P has the potential to overcome the scalability problem associated with traditional client-server-based CDNs, it brings a set of new challenges. First, because uplink capacity of typical broadband access technologies is limited when compared to downlink, only a small subset of participating peers may be able to become contributing sources when bandwidth-intensive content is distributed. Second, since video streaming allows discarding of video frames anytime after their playback, the availability of video frames in user buffers for access by others becomes time-bounded and time-changing such that it aggravates the challenge associated with content discovery on P2P networks. Third, due to uncertainty in the behavior of peers sourcing video, users may experience excessive delays in video reception. As video streaming requires an orderly and timely delivery of video frames for a smooth playback, a video distribution scheme which minimizes the impact of distribution source losses on P2P networks is desired.

To achieve these goals, the authors proposed the design of a new video distribution network model and accompanying video segment discovery and reception schemes with QoS support. The following contributions were made through this research:

- Design of a new streaming video distribution network model called Virtual Theater Network.
- Design of a segmented video stream reception scheme and accompanying scheduling algorithm for orderly and timely video segment retrievals. It enables users with limited transmit bandwidth (i.e., transmit bandwidth << the nominal streaming rate) to become contributing sources.
- Design of an advertisement scheme for the discovery of available video segments in user buffers which incorporates the parameters of the video reception scheduling algorithm. It greatly simplifies the discovery process such that one advertisement and one query are sufficient to post and retrieve the lifetime video segment availability of a user.
- QoS support in mitigating the video viewing interruptions in the face of excessive delays, including ones caused by multiple video distribution source losses.

Related Work

Previous studies on P2P-based video distribution networks can be classified into two categories depending on the number of distribution sources from which participating users may request video streams. Under the single distribution source model, a requesting peer receives the entire video stream from one peer in the network. Depending on the schemes, peers self-organize themselves in a logical topology either in the form of chains [1], loops [2] or trees [3], [4]. The aim of these networks is to support live-media applications such as news tickers and real-time stock updates, which distribute low-bandwidth contents to many users. These schemes fall short in the support of high-bandwidth stored streaming video distribution services because they have no consideration for the asymmetric bandwidth availability of most of the widely used broadband access services available to consumers.

Under the multiple distribution source model, a requesting peer receives video streams from multiple peers. The combinations or concatenations of all streams reconstruct the original video. Some schemes split the video into multiple decodable layers, each of which is transmitted from a different source [5], [6]. While this approach avoids total loss of service in times of network failures and source losses, it incurs a large overhead to support the resiliency design. Other schemes divide the total length of the video in time and a sequence of video segments is transmitted from different sources [7-9]. The scheme presented here belongs to this category. The major difference between this scheme and those proposed by other authors lies in the assumption of how long the received video segments will remain in the user system once they are played back. This scheme assumes that they are discarded after a certain period of time and that their availability for retrieval by other peers is time bounded. Other schemes assume, implicitly or explicitly, long-term availability of downloaded content at the users' permanent storage system.

Architecture

Virtual Theater Network is a network model that aims to enable large-scale, on-demand, peer-to-peer stored-video streaming service over the Internet, which incorporates a hybrid architecture between client-server and peer-to-peer computing. Central to this model is a set of Virtual Theaters which provide a means to mass distribute video streams to communities of users. Within each Virtual Theater there exists a content distributor, known as a VT Distributor. A VT Distributor manages one or more VT Rooms in order to service the video distribution needs of users. A VT Room is a group of peers that forms a P2P community to receive and distribute a video stream.

Figure 1 illustrates the Virtual Theater Network model. There are multiple instances of Virtual Theaters throughout the Internet and this is depicted in the figure as Virtual Theaters 1, 2 and N. In this example, Virtual Theater 1 consists of VT Distributor 1 and three VT Rooms: VT Rooms 1, 2 and M. Each VT Room distributes a different video title and is created when the first user begins receiving the video feed from the VT Distributor. Subsequent users desiring to watch the same title of the video join the respective VT Rooms. As they join, they discover other peers in the room. Small circles within each VT Room in the figure represent the peers that joined the VT Room. The video in the VT Room is divided into a time sequence of small segments. The peers discover them in the buffers of other peers and retrieve them in their playback order. As video segments are being downloaded, the receiving peer makes them available for others to retrieve.

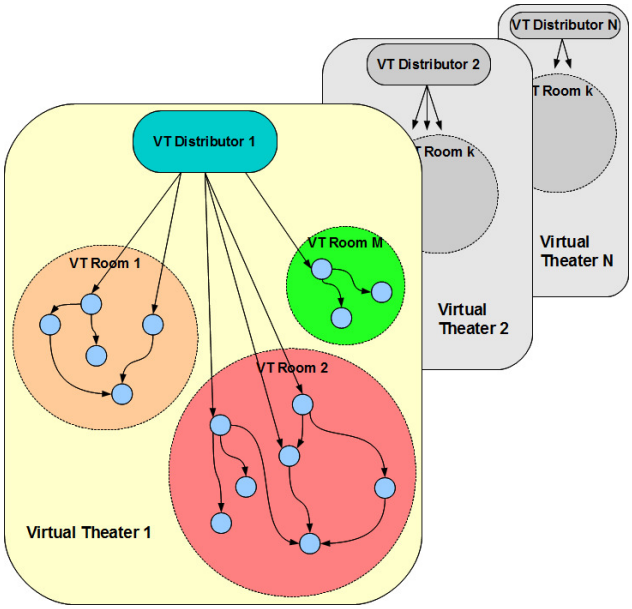


Figure 1. A Conceptual View of Virtual Theater Network

The main challenge in the design of video distribution service based on the proposed architecture is threefold: 1) how to manage the orderly and timely delivery of video segments that contribute to the self-sustainability of a VT Room, 2) how to organize dynamically changing video segment availability information within a VT Room to provide effective video segment advertisement and discovery service, and 3) how to mitigate the video viewing interruptions in face of excessive delays due to video distribution source losses. These issues are addressed in the following sections.

Video Segment Reception Management

This section describes Sliding Batch, a video segment reception scheme used in a VT Room and defines key concepts. The video segment reception scheme in Sliding Batch is expressed in terms of segments, epochs and batches. The definition of each is given below.

Segments

A video stream is a continuous flow of a sequence of compressed video frames transmitted over a network so that the recipient may play back the video frames as they arrive. In Sliding Batch, a block of a sequence of video frames makes up a Segment, S_i, and the concatenation of segments in their correct time sequence creates a video, V. Segments are similar in concept to chapters in a DVD and may vary in size and length. The number and size of segments in a video are VT-Room specific parameters. A segment is the basic unit of video exchanges among the users of a VT Room.

A segment, S_i, is characterized by its sequence position, i, in V, a set of frames, f_k, that belongs to S_i, its starting playback time, $\alpha(S_i)$, and the batch it belongs to, $\beta(e_j)$

$$V = \{S_i, i = 1,2,\cdots,N\}$$

$$S_i = \begin{cases} \{f_k, k = 1,2,\cdots,n_i, \quad n_i \leq F\} & if\ i = 1 \\ \{f_k, k = n_{i-1}+1, n_{i-1}+2,\cdots,n_i, \quad n_i \leq F\} & \\ \qquad if\ 2 \leq i \leq N & \end{cases}$$

$$S_i \cap S_j = \emptyset$$

$$\alpha(S_i) = \begin{cases} t_0 & if\ i = 1 \\ \alpha(S_{i-1}) + \delta(S_{i-1}) & if\ 2 \leq i \leq N \end{cases}$$

$$S_i \in \beta(e_j)$$

where N is the total number of segments in V, F is the last frame number in V, t_0 is the time the user joined the VT Room and began playing back the first segment, S_i, and $\delta(S_i)$ is the playback duration of S_i. The description of $\beta(e_j)$ is included later in this discussion.

Let $|V|$ and $|S_i|$ be the size of V and S_i, respectively. Let $\delta(V)$ be the total video playback time. Then, η, the nominal streaming rate of a video is given by $\eta = |V| / \delta(V)$, where $|V| = \sum_{i=1}^{N} |S_i|$. $\sum_{i=1}^{N}$ Accordingly, the playback duration of S_i, $\delta(S_i)$, is defined as $\delta(S_i) = |\delta(S_i)| / \eta$.

Epochs

In Sliding Batch, the lifetime of a video stream is divided into a sequence of time intervals, known as epochs. There are N epochs in a V and their duration may vary from epoch to epoch. Both the number and duration of epochs in a video are VT-Room-specific parameters. An epoch, e_i, is characterized by its starting epoch time, $\alpha(e_i)$, its duration, $\delta(e_i)$ and its associated batch, $\beta(e_i)$. An epoch is closely related to the playback property of a segment and described by Equations (1) – (3).

$$\alpha(e_i) = \alpha(S_i) \tag{1}$$

$$\alpha(e_{i+1}) = \alpha(e_i) + \delta(e_i) \tag{2}$$

$$\delta(e_i) = \frac{|S_i|}{\eta} = \delta(S_i) \tag{3}$$

Batches

A batch, $\beta(e_i)$—each of which is associated with an epoch—is a set of segments whose downloads are initiated simultaneously at the beginning of (e_i). There are total of N batches in a video and each batch consists of a set of segments unique to itself, except for those batches with an empty set of segments. $\beta(e_i)$ is characterized by an associated epoch, e_i, a set of video segments, and a set of streaming sessions that are initiated at epoch e_i, each with rate r_j; refer to Equation (4).

$$\beta(e_i) = \begin{cases} \{S_j, j = 1,2,\cdots,n_j, \quad n_j \leq N\} & if\ i = 1 \\ \{S_j, j = n_{j-1}+1, n_{j-1}+2,\cdots,n_j, \quad n_j \leq N\} & \\ \qquad if\ 2 \leq i \leq N\ and\ n_{j-1} < N & \\ \emptyset \qquad otherwise & \end{cases} \tag{4}$$

$$\beta(e_i) \cap \beta(e_k) = \emptyset$$

$$A(S_j) = \alpha(e_i), \qquad \forall S_j, \qquad S_j \in \beta(e_i)$$

where $A(S_j)$ is the starting download time of S_j.

The ending download time of segment j, $\Omega(S_j)$, differs from segment to segment and is the ending playback time of segment j; refer to Equation (5).

$$\Omega(S_j) = A(S_j) + \frac{|S_j|}{r_j}, \quad S_j \in \beta(e_i) \tag{5}$$

Figure 2 illustrates the relationship between segments, epochs and batches in a simplified video reception scenario. In this example, a video is divided into four segments (N = 4) of varying lengths. Batch $\beta(e_1)$ consists of segments S_1 and S_2. Segment downloading for $\beta(e_1)$ was initiated at time

$\alpha(e_1)$ for both S_1 and S_2 at rates r_1 and r_2 , respectively. Batch $\beta(e_2)$ consists of segments S_3 and S_4. Segment downloading for $\beta(e_2)$ was initiated at time $\alpha(e_2)$ at rates r_3 and r_4 , respectively. No segment was associated with $\beta(e_3)$ or $\beta(e_4)$.

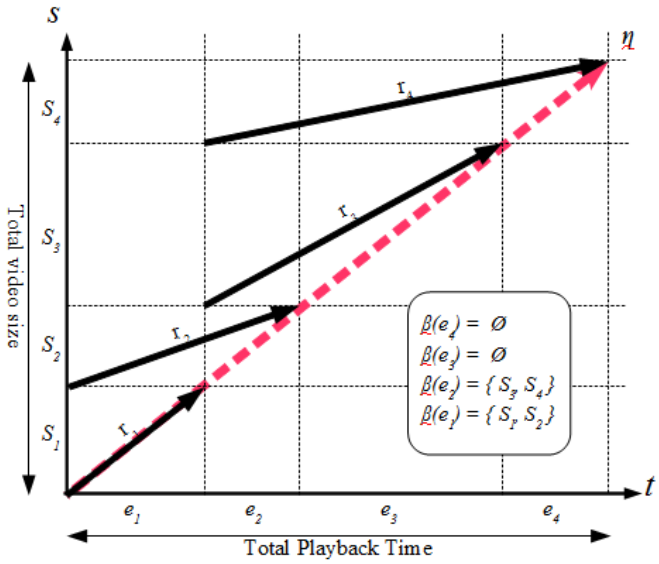


Figure 2. Relationship Among Segments, Epochs and Batches

Batch Size Determination

An important parameter of Sliding Batch is the number of segments in a batch, or the size of a batch, $|\beta(e_i)|$. It determines the beginning downloading time and the rate of each segment in the batch. The batch size, $|\beta(e_i)|$, is the total number of segments that a user begins downloading simultaneously at time $\alpha(e_i)$ and was determined by each user's available receive bandwidth and buffer space; refer to Equation (6)

$$|\beta(e_i)| = min\left(\overline{N(e_i)},\quad |\beta_R(e_i)|,\quad |\beta_B(e_i)|\right) \quad (6)$$

where $\overline{N(e_i)}$ is the number of remaining segments yet to be downloaded at time $\alpha(e_i)$, $|\beta_R(e_i)|$ is the rate-limited batch size at time $\alpha(e_i)$ and $|\beta_B(e_i)|$ is the buffer-limited batch size at time $\alpha(e_i)$.

The number of remaining segments, $\overline{N(e_i)}$, is defined in Equation (7).

$$\overline{N(e_i)} = \begin{cases} N & if\ i = 1 \\ N - \sum_{k=1}^{i-1} |\beta(e_k)| & if\ 2 \le i \le N \end{cases} \quad (7)$$

The rate-limited batch size, $|\beta_R(e_i)|$, refers to the size of a batch being computed solely on the available receive bandwidth, R_A, of the user and is determined by the maximum number of concurrent segment downloading sessions that can be sustained at time $\alpha(e_i)$; refer to Equation (8)

$$|\beta_R(e_i)| = \begin{cases} m_1, \\ \quad \exists max(m_1) | \sum_{k=1}^{m_1} r_k \le R_T,\ m_i \le N \\ \quad\quad if\ i = 1 \\ m_i - m_{i-1}, \\ \quad \exists max(m_i) | \sum_{k=m_{i-1}+1}^{m_i} r_k \le R_A(e_i), \\ \quad\quad m_i \le N \quad if\ 2 \le i \le N \end{cases} \quad (8)$$

where $R_A(e_i)$ is the receive bandwidth available at time $\alpha(e_i)$ and $R_T = R_A(e_1)$.

Similarly, the buffer-limited batch size, $|\beta_B(e_i)|$, was computed solely on the available buffer size, B_A, as if there were an infinite amount of receive bandwidth available. $|\beta_B(e_i)|$ is determined by the maximum number of concurrent segment downloads that can be sustained at time $\alpha(e_i)$; refer to Equation (9)

$$|\beta_B(e_i)| = \begin{cases} m_1, \\ \quad \exists max(m_1) | \sum_{k=1}^{m_1} |S_k| \le B_A(e_1), \\ \quad\quad m_1 \le N \quad if\ i = 1 \\ m_i - m_{i-1}, \\ \quad \exists max(m_i) | \sum_{k=m_{i-1}+1}^{m_i} |S_k| \le B_A(e_i), \\ \quad\quad m_i \le N \quad if\ 2 \le i \le N \end{cases} \quad (9)$$

where $B_A(e_i)$ is the available buffer size at time $\alpha(e_i)$.

The details of receive bandwidth and buffer management in relation to the scheduling algorithm are given in a study by Okuda and Znati [10] that includes a description of Restrained Sliding Batch, a variant of Sliding Batch that tames the aggressive segment pre-fetch behavior of the original scheme.

Figure 3 illustrates an example of how a user may receive segments in batches. In this example, a streamed video consists of 24 equally sized segments. The total receive bandwidth, R_T, of the user is twice the nominal streaming rate of the video segment. The total download buffer space, B_D, can accommodate a maximum of 10 simultaneous segment downloads. Fifteen batches are needed to initiate downloading of all segments. Notice that the first two batches, $\beta(e_1)$ and $\beta(e_2)$, are rate limited while the next eight batches, $\beta(e_3)$ through $\beta(e_{15})$, are buffer limited.

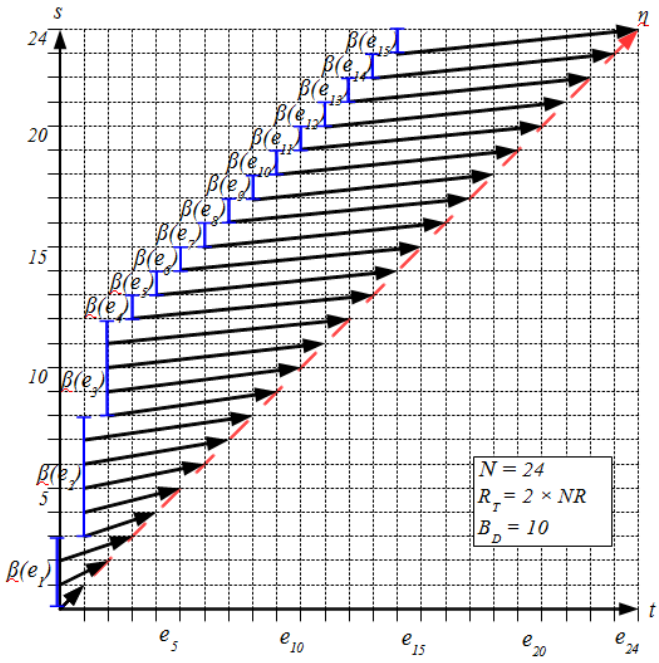


Figure 3. Sample Segment Receptions: Bandwidth and Buffer Limited Case

User Profile and VT Room Profile

The parameters used in Sliding Batch for the computation of batch size belong to either the VT Room profile or User profile. The VT Room profile describes the attributes of a video being distributed in a VT Room. These include parameters such as the size of a streaming video, $|V|$, playback duration, $\delta(V)$, the total number of segments, N, and the size of each segment, $|S_i|$. The VT Room profile is given to all users in each VT Room at the time they join the VT Distributor.

User profile describes the attributes of a user, primarily its resource availability, and consists of the following parameters: the time the user joined the VT Room, t_0, the total receive bandwidth set aside for the streaming service, R_T, the downloading buffer size, $|B_D|$, and the size of post-playback buffer space, $|B_H|$. $|B_H|$ determines how long a segment will remain in the user buffer after its playback. Users in a VT Room advertise their user profile through the advertisement and discovery scheme described in the next section.

Video Segment Advertisement and Discovery

This section describes Virtual Chaining, the video segment advertisement and discovery scheme used in a VT Room. Virtual Chaining allows users to cooperatively maintain a collection of user profiles, known as a state table, to share their segment reception state information with other users.

State Table

A state table is a collection of user profiles maintained cooperatively among the members of a VT Room. It describes each user's segment reception state and the transmit bandwidth availability. An entry in the state table consists of the following fields: IP address of the user advertising its state, parameters of the user profile (t_0, R_T, $|B_B|$, $|B_H|$), available transmit bandwidth (T_A), and the time of its entry. This is depicted in Figure 4.

IP	t_0	R_T	$\|B_D\|$	$\|B_H\|$	T_A	Time of Entry

Figure 4. Entry Fields of a State Table

An entry is added to the state table when a new user joins a VT Room. It is removed when the last video segment is dropped from the user's post-playback buffer. This condition can be determined by comparing the current time against when the user joined the VT Room plus the video playback duration and how long the video segments stay in the post-playback buffer; refer to Equation (10).

$$t_c > t_0 + \delta(V) + \frac{|B_H|}{\eta} \qquad (10)$$

State Table Sharing

The state table is shared among the users of a VT Room in the following manner. The VT Distributor maintains the tail-end portion of the state table, which contains user profiles of the last n users who joined the VT Room. A newly arrived user, U_i, receives the state table from the VT Distributor and reports its profile. The VT Distributor adds U_i's profile in the state table and drops the oldest entry if the table becomes greater than n. The VT Distributor waits for the next user arrival. In the meantime, U_i examines the received state table and tries to identify other users who may be able to provide segment distributions. If more users need to be discovered, U_i requests and maintains U_{i-n}, the oldest entry in the state table received from the VT Distributor. U_{i-n}'s state table consists of the user profiles of U_{i-n} through U_{i-2n} and potentially beyond if U_{i-n} had requested the state table from U_{i-2n}. This process is repeated until a qualified distribution source is located. If no qualified distribution source is found after stepping through the chain of state tables, U_i requests the direct video feed from the VT Distributor.

Figure 5 illustrates a sample trace of the state table sharing process. In this figure, circles represent users and shaded boxes below each circle represent portions of the state table maintained by each user. While all users maintain a portion of an overlapping state table, this figure only shows the ones maintained by users whose user ID is a multiple of n. Each shaded box contains n entries of user profiles. The user pointed to by the oldest entry in the state table is denoted by the dashed arrow extending from the box to the appropriate user. Solid arrows represent the transfer of state table entries.

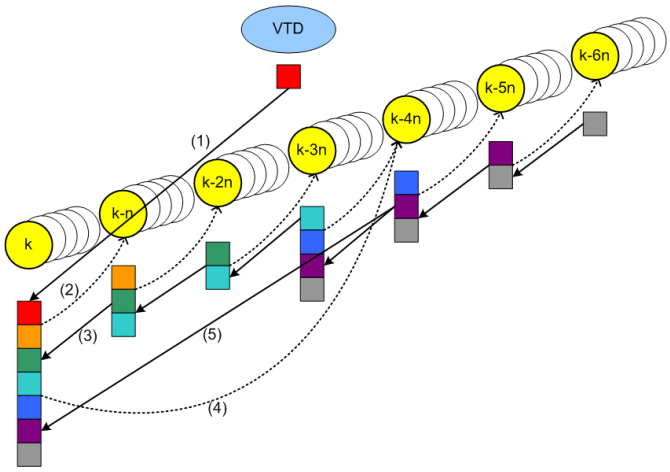


Figure 5. A Sample View of State Table Sharing Instances

A newly arrived user, U_k, joins a VT Room and receives the state table from the VT Distributor (step 1), which contains the user profiles of U_{k-1} through U_{k-n}. To discover more users in the VT Room, U_k requests older state table entries from U_{k-n} (step 2). U_{k-n} maintains the user profiles of U_{k-n-1} through U_{k-2n} it received from the VT Distributor at the time it joined. In this example, U_{k-n} also has the user profiles of U_{k-2n-1} through U_{k-4n}, which was received from U_{k-2n}. All of these entries are sent from U_{k-n} to U_k (step 3). To further discover older users, U_k requests U_{k-4n} to send its portion of the state table (step 4). U_{k-4n} sends the user profiles of U_{k-4n-1} through U_{k-7n} to U_k (step 5).

If U_k does not receive a response from the user (e.g., U_{k-n}) from which it requested an older state table, the next oldest entry in the state table (i.e., U_{k-n+1}) will be contacted.

Due to its simple operation, Virtual Chaining is relatively easy to implement, deploy, and study its behavior. A distributed and redundant state table, available at participating users, offers resiliency such that a loss of a few users does not break the segment advertisement, discovery or distribution operation. Virtual Chaining is fair, in terms of the carried workload among the users, such that no single user is expected to perform more work than any other. Virtual Chaining is also scalable in that the workload placed upon each user remains a constant regardless of the size of the membership in the P2P community.

QoS Support

Streaming applications that operate over a network with fluctuating traffic delays, such as the Internet, employ a playout buffer, B_P. The goal of the playout buffer is to prevent video frame starvation (i.e., absence of video frames in a buffer) during playback. This is achieved by pre-fetching an initial portion of a video stream and withholding its playback for a predetermined duration of time. The delay incurred by this operation is referred to as playout delay, D_P, or start-up delay. In exchange for inducing delay, it is hoped that the subsequent delays during the lifetime of video playback may be absorbed by the playout buffer. Playout delay should be long enough to cope with typical delays seen on the Internet, yet short enough for users to tolerate the initial waiting time.

The challenge to incorporating the traditional network delay coping mechanism in Sliding Batch is how to deal with the loss of distribution sources. To address this issue, a set of delay management mechanisms was proposed. Extended Playout Delay (EPD) and Expedited Segment Reception (ESR) work together to prevent buffer under-run conditions during multiple instances of distribution source losses.

Extended Playout Delay

Extended Playout Delay (EPD) is designed to prevent a buffer starvation condition after an excessive delay is detected. This is achieved by having users wait extra time before the initial video playback can begin. It differs from the traditional delay coping mechanism in that EPD introduces the concept of level of protection (n), which aims to separate the excessive delay detection period, (D_P), from the initial playout delay period (i.e., extended playout delay, D_E). Equation (11) depicts the relationship between n, D_P and D_E

$$D_E = n \cdot D_P = \frac{n \cdot |B_P|}{\eta} \tag{11}$$

where $|B_P|$ is the size of playout buffer. n corresponds to the number of times a user may encounter excessive delays (i.e., the number of distribution source losses) and not experience video presentation interruptions.

Figure 6 displays a logical view of a sample extended playout buffer, B_E, with an excessive delay protection level of $n = 4$.

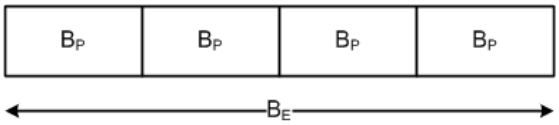


Figure 6. A Sample Extended Playout Buffer

Let D_{En} be the start-up latency introduced by EPD to provide an n^{th} level of excessive delay protection. At D_{E1} = 1 x D_P, EPD only provides delay absorptions up to D_P with no protection against a distribution source loss, just as the traditional playout buffer does. At D_{E2} = 2 x D_P, EPD offers a one-time distribution source loss protection during the lifetime of a segment download, in addition to delay absorptions up to D_P. It provides an uninterrupted video presentation if a user loses one distribution source and begins receiving the segment from a new distribution source. However, if the user loses the new distribution source, there will be a playback interruption. At D_{E3} = 3 x D_P, a two-time distribution source losses can be tolerated during the lifetime of a segment download, in addition to delay absorptions up to D_P. By extending the start-up delay, the level of protection against multiple instances of distribution source losses over the lifetime of a segment download can be improved.

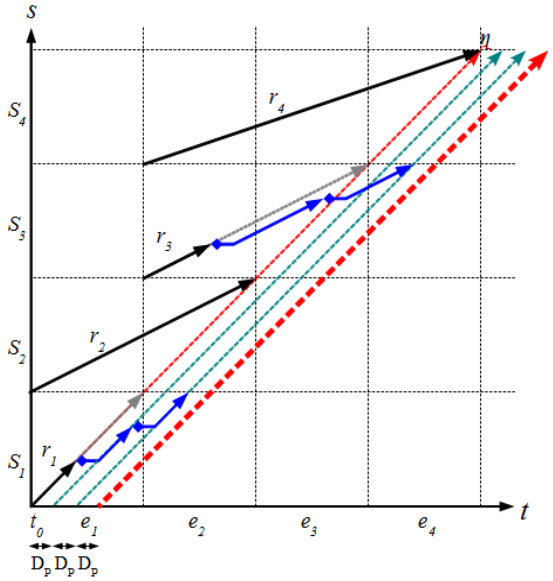


Figure 7. Segment Receptions with Extended Playout Delay

A sample segment reception with EPD is depicted in Figure 7. In this example, start-up delay has been extended to D_{E3} = 3 x D_P. The dotted arrow lines running diagonally across the middle of the figure represent the nominal streaming rates and show the playback positions of the streamed video at three levels of protections. Segments S_1 and S_3 both experience an excessive delay twice during their download, which is depicted by horizontal lines with diamond-shaped starting points. Once the excessive delay condition is declared, the user begins retrieving the affected segment from another distribution source. Note that, even after moving to a new distribution source twice, a sufficient amount of data has been pre-fetched in the extended playout buffer to provide a delay absorption up to D_P.

Expedited Segment Reception

Expedited Segment Reception (ESR) offers protection against a distribution source loss by increasing the segment downloading rate. By expediting the ending time of a segment download, ESR attempts to gain sufficient time to recover from delays experienced during segment receptions.

Let r_i^{En} be the rate of ESR for downloading segment S_i at n^{th} level of protection against excessive delays. At $r_i^{E1} = |S_i| / (\Delta(S_i) - D_P)$, given $\Delta(S_i) > D_P$, where $\Delta(S_i)$ is the downloading duration of S_i (i.e., $\Delta(S_i) = \Omega(S_i) - A(S_i)$), one-time distribution source loss protection can be achieved during the lifetime of an S_i download. It assures an uninterrupted video viewing experience by the user even if a distribution source is lost and the delayed segment is retrieved from a new distribution source. However, if the user loses the distribution source again, there will be a playback interruption. At $r_i^{E2} = |S_i| / (\Delta(S_i) - 2D_P)$, given $\Delta(S_i) > 2D_P$, two-time distribution source loss protection can be offered during the lifetime of an S_i download. At $r_i^{E3} = |S_i| / (\Delta(S_i) - 3D_P)$, given $\Delta(S_i) > 3D_P$, a three-time distribution source loss protection can be offered during the lifetime of S_i download. A higher degree of protection can be achieved with a relatively small amount of increase in the rate of segment download.

Figure 8 shows an example of how video segments may be received under ESR. In this example, S_1 is protected against one-time distribution source loss by receiving the segment at r_1^{E1}. S_2 and S_3 are protected against two-time source losses by increasing the reception rate to r_i^{E2}. At r_i^{E3}, S_4 can withstand three-time distribution source losses. The example shows that excessive delays have been observed while downloading segments S_1 and S_3, but they did not cause video playback interruptions because a sufficient amount of data was pre-fetched through ESR.

Downloading of each segment can be associated with a different degree of protection through ESR. For example, if a measurement shows that a significantly higher rate of distribution source loss is experienced in downloading S_i, a higher level of protection can be afforded to the reception of S_i. Let L_i be the maximum number of distribution source losses a user anticipates when downloading an S_i segment.

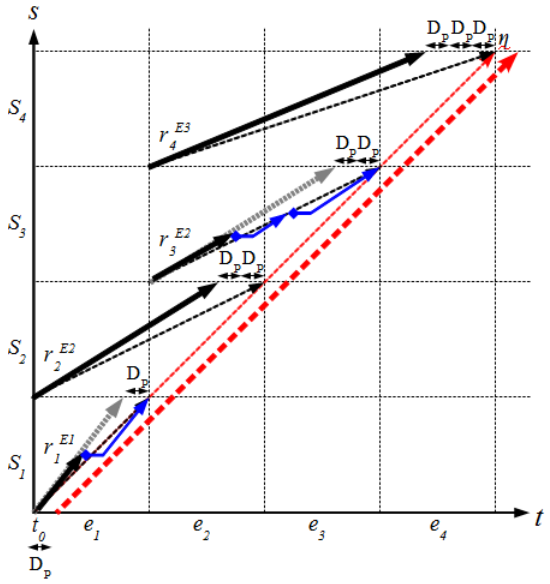


Figure 8. Sample Segment Receptions with Expedited Segment Reception

The expedited rate of reception, r_i^{EL}, that protects against L number of distribution losses is given by Equation (12).

$$r_i^{EL} = \frac{|S_i|}{\Delta(S_i) - L_i \times D_P} \qquad (12)$$

In order to determine the level of protection needed for each segment download, the statistic on the loss of distribution sources must be collected. This is achieved by informing the VT Distributor every time a user experiences a distribution source loss. The VT Distributor keeps the statistic of distribution source losses for each segment and shares it with users as they join the VT Room.

Let P_i be the probability of a user experiencing a distribution source loss when downloading S_i, assuming loss of distribution sources are IID. To achieve a successful download of S_i at or above a protection goal, g such as $g = 0.95$, the following condition must be met:

$$g \geq 1 - P_i^{L_i}$$

The level of protection needed for an S_i download can be computed by solving for L_i.

The user executes the Excessive Delay Protection Algorithm, as depicted in Table 1, to determine the amount of Extended Playout Delay and the rate of Expedited Segment Reception. The strategy used in this algorithm is to let the user wait as long as it is willing at the initial playback time through EPD. If necessary, expedite individual segment receptions through ESR can be expedited. Let L be the maximum level of protection required to meet g. Initially, L is set to the maximum value of L_i, the greatest level of protection required among all segment receptions. Let w be the maximum time a user is willing to wait for the start-up latency. If the extended playout delay, $D_E = L \times D_P$, is greater than w, L is set to $[w/D_P]$. This is the level of protection offered by EPD. For each segment that belongs to a batch, $\beta(e_i)$, a need for an additional level of protection through ESR is investigated. If the level of protection, L_i, required for downloading S_i is greater than the level of protection provided by the EPD (L), the rate at which the segment is downloaded will be increased to $r_i^E = |S_i| / (\Delta(S_i) - (L_i - L) \times D_P)$. If not enough receive bandwidth is available, the segment download will not be initiated.

Table 1. Excessive Delay Protection Algorithm

```
1:   // initialize
2:   L = max(L_i)
3:   // let user wait as long as it is willing
4:   if (L x D_P > w) then
5:   L = [w/D_P]
6:   endif

7:   for each segment ∈ β(e_i) d ∈ o
8:   // increase the download rate as needed
9:   if (L_i > L) then
10:  r_i^E = |S_i| / (Δ(S_i) – (L_i – L) x D_P)
11:  endif
12:  // not enough RxBW to meet the protection
       requirement
13:  if (r_i^E > R_A) then
14:  print warning and break
15:  endif
16:  endif
```

Simulation Design and Analysis

This section describes the design and analysis of experiments performed on a Virtual Theater Network. A software model was created to simulate the behavior of a VT Room. Two types of experiment was conducted. The focus of the first type of experiments is to study how well the proposed video distribution scheme would alleviate the load on the VT Distributor under different operating environments. The results of the first set of simulation studies are available from Okuda and Znati [10]. The second set of experiments focuses on how well the proposed QoS scheme would mitigate the impact of distribution source losses on the video presentation. The description of the model, the design of the experiments, and the analysis of simulation study are given below.

Description of the Model

The simulated VT Room consists of a VT Distributor and a series of user processes that arrive at the VT Room. The VT Distributor supplies the parameters of the VT Room profile to the newly joining users, such as the total video playback time (120 minutes), the nominal streaming rate (1.0 Mbps) and the total number of segments (24) in the video. Note that the values in parentheses denote the default values used in the experiments. For simplicity of simulation for this study, the video was divided into equal segment lengths and equal playback times (5 minutes).

The user processes simulate the behavior of peers joining the VT Room, discovering other users, identifying possible distribution sources, receiving video segments, distributing video segments as requests arrive, and departing from the VT Room. The inter-arrival time of user processes is exponentially distributed (a mean of 10 seconds). To reflect the asymmetrical nature of the transmit and receive bandwidth capacity of typical broadband access technologies, each user was equipped with a fixed receive bandwidth (2.0 Mbps) and varying transmit bandwidth (30% to 100% of the receive bandwidth; uniformly distributed). Each user executes Virtual Chaining to identify possible distribution sources and implements Restrained Sliding Batch to receive video segments. All experiments simulate the bandwidth-limited network environment where a sufficient amount of download buffer exists at each user ($|B_H| \geq N$). The default post-playback buffer size allows a segment to remain in buffer for a finite period of time (15 minutes) after its playback.

Experimental Design and Analysis

Two sets of experiments were designed to study the effectiveness of QoS schemes in mitigating video presentation interruptions when users experience excessive delays while receiving video segments. Each experiment was measured against Chaining [1].

The first set of experiments studied the effects of the size of the extended playout delay in reducing the video presentation interruptions under different rates of premature user departures from the VT Room. Extended Playout Delays were varied from 1 x D_P to 4 x D_P. The probability of premature node departure, P, was varied from 0.1 to 0.4. The time a node may spend before prematurely departing from the VT Room was uniformly distributed during the playback of the entire video. The total number of video presentation interruptions experienced by participating users was normalized to the total number of distribution source losses being detected in the VT Room.

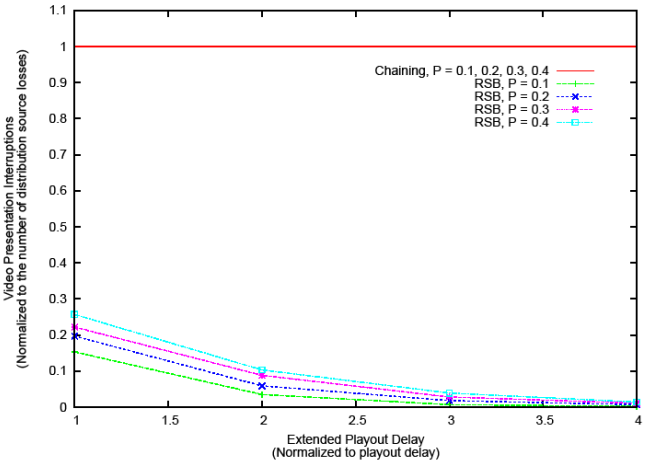


Figure 9. Effects of Extended Playout Delay – I

In Chaining, all distribution source losses being detected by the users resulted in the video presentation interruptions, as they use a traditional playout buffer mechanism, regardless of the rate of user departures from the VT Room. This is depicted in Figure 9 by the horizontal line drawn at the 1.0 mark. In Restrained Sliding Batch, through the implementation of the proposed QoS scheme, not all distribution source losses being detected result in the interruption of the video presentation. The user may experience interruptions only if a segment reception encounters a greater number of distribution source losses than all other segments in the same batch (i.e., maximum distribution source losses of a batch, $L_{\beta(ei)}$) Furthermore, a video presentation interruption can occur only if the sum of the maximum distribution source losses of all batches results in an accumulated delay beyond the extended playout delay. Let Γ be the total number of video presentation interruptions being experienced by a user for the duration of the video playback. Γ is defined as:

$$\Gamma = \left\lceil \frac{\sum_{i}^{N} L_{\beta(ei)} \times \text{playout delay}}{\text{extended playout delay}} \right\rceil$$

In Restrained Sliding Batch, at 1 x D_P, 15% to 26% of distribution source losses being detected resulted in actual video presentation interruptions, when 10% to 40% of the nodes prematurely departed from the VT Room. At the extended playout delay of 2 x D_P, the rate of video presentation interruptions decreased between 4% and 10% when P was varied from 0.1 to 0.4. When the size of the extended playout buffer was increased to 4 x D_P, less than 1% of all distribution source losses being detected by users resulted in actual video presentation interruptions when 10% of the total nodes prematurely leave the VT Room. At a rate of 40% of premature node departure, roughly 1% of the de-

tected source losses resulted in video presentation interruptions.

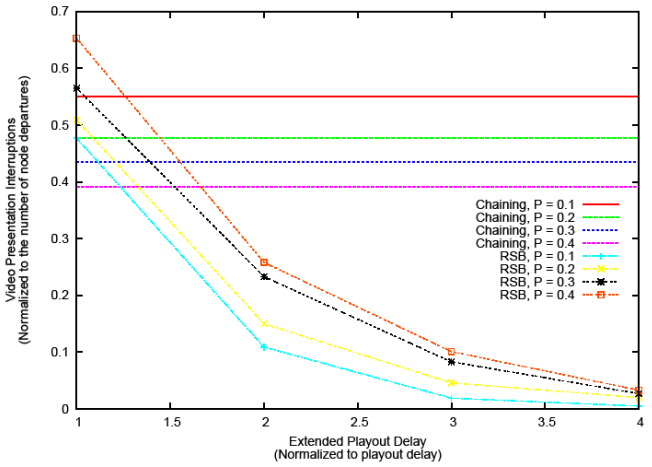


Figure 10. Effects of Extended Playout Delay – II

Another way to express the effects of extended playout delay on the video presentation interruptions under different node departure rates is shown in Figure 10. In this figure, the total number of video presentation interruptions is normalized to the total number of premature node departures under the same settings, as shown in the previous set of experiments. Note, in Restrained Sliding Batch, a premature user departure may result in multiple instances of distribution source losses experienced by other users. Figure 10 shows how many instances of video presentation interruptions are introduced when one node departs prematurely from the service. In Restrained Sliding Batch, every premature node departure resulted in a video presentation interruption at the rate of 48% to 65% at 1 x D_P, when P is varied from 0.1 to 0.4. When the size of the extended playout buffer was doubled, the interruption rates halved at P = 0.4 and quartered at P = 0.1. When the extended playout delay was at 4 x D_P, less than 1% of premature node departure resulted in a video presentation interruption when P = 0.1 and roughly 3% of premature node departures resulted in video interruptions when 40% of users left the service prematurely.

In Chaining, every premature node departure resulted in a video presentation interruption at rates between 55% and 39% when P was varied from 0.1 to 0.4. At first glance, the simulation result seems counter intuitive in that the rate of video interruptions decreased as the rate of node departures increased. This is because, in Chaining, as the rate of node departure increases, a large percentage of users begin relying on the central server for video distribution feeds rather than their peers. The next set of experiments proves this point.

The second set of experiments studied the effects of premature node departure rates on the load on the VT Distributor under varied user arrival rates to the VT Room. The results are shown in Figure 11. The peak transmit bandwidth demand on the VT Distributor was normalized to the peak transmit bandwidth demand on a traditional client-server-based video distribution network. At P = 0, no node prematurely departs from the VT Room and the results from this simulation were used as reference. For $0.1 \leq P \leq 0.4$, both Chaining and Restrained Sliding Batch have similar rates of load increase on the VT Distributor as the rate of premature node departure increased. The main difference between the two schemes is that Restrained Sliding Batch with QoS extension requires only a third or less of resources from the central server when compared to Chaining.

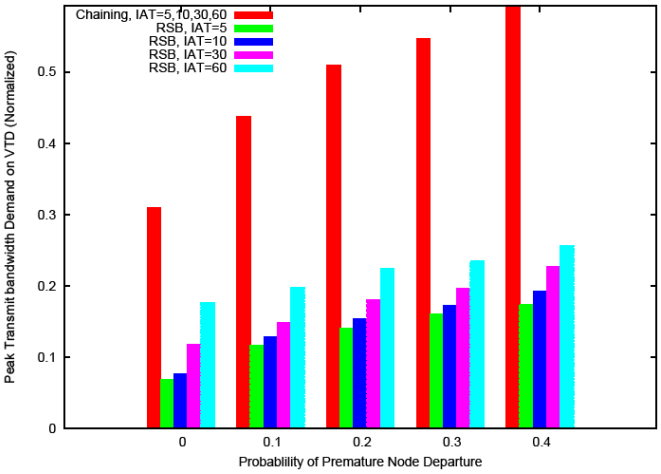


Figure 11. Effects of Node Departure Rate

These two sets of experiments verified the effectiveness of QoS schemes in reducing the number of video presentation interruptions, when users depart from the network prematurely. A linear increase in the size of the playout buffer resulted in logarithmic decreases in the rate of video presentation interruptions. Furthermore, relatively small increases in the VT Distributor load was observed when the probability of premature node departure was raised.

Conclusion

The authors proposed a design for a new streaming video distribution network model called Virtual Theater Network, which allows organization of peer-to-peer communities to support the distribution of videos among the community members. The model employs a segmented video stream reception scheme with its accompanying scheduling algorithm for orderly and timely video segment retrievals that allow contributions from users with limited transmit band-

width availability. QoS extension of the distribution scheme allows reduction in the number of video presentation interruptions when excessive delays are observed. The model also employs a video segment availability advertisement and discovery scheme, which incorporates the parameters of the scheduling algorithm. It enabled the advertisement and query of dynamically changing segment availability information of each user in one advertisement and one query. The simulation study showed dramatic improvements in the video presentation interruption occurrences under the proposed QoS scheme when distribution sources depart prematurely from the network.

References

[1] Sheu, S. & Hua, K. A. (1997). Virtual Batching: A New Scheduling Technique for Video-on-Demand Servers. *Proceeding of the International Conference on Database Systems for Advanced Applications*, (pp. 481-490).

[2] Kusmierek, E., Dong, Y. & Du, D. (2005). Loopback: Exploiting Collaborative Caches for Large-Scale Streaming. *Proceedings of ACM/SPIE MMCN*.

[3] Banerjee, S., Bhattacharjee, B. & Kommareddy, C. (2002). Scalable application layer multicast. *Tech. Rep., UMIACS TR-2002*.

[4] Tran, D., Hua, K. & Do, T. (2003). Zigzag: An efficient peer-to-peer scheme for media streaming. *Proceedings of IEEE INFOCOM*.

[5] Padmanabhan, V., Wang, H., Chou, P. & Sripanidkulchai, K. (2002). Distributing streaming media content using cooperative networking. *Proceedings of ACM/IEEE NOSSDAV*.

[6] Rejaie, R. & Ortega, A. (2003). PALS: Peer-to-Peer Adaptive Layered Streaming. *Proceedings of ACM NOSSDAV*.

[7] Hefeeda, M, Bhargava, B. K. & Yau, D. K. (2004). A hybrid architecture for cost-effective on-demand media streaming. *IEEE Computer Networks*, 44(3).

[8] Shan, Y. & Kalyanaraman, S. (2003). Hybrid video downloading/streaming over peer-to-peer networks. *Proceedings of IEEE ICME*.

[9] Hefeeda, M., Habib, A., Botev, B., Xu, D. & Bhargava, B. (2003). PROMISE: Peer-to-Peer Media Streaming Using CollectCast. *Proceedings of ACM Multimedia*.

[10] Okuda, M. & Znati, T. (2006). Virtual Theater Network: Enabling Large-Scale Peer-to-Peer Streaming Service. *Proceedings of Distributed Multimedia Systems*.

Biography

MASARU OKUDA received the B.S. degree in Information System and Computer Science from Brigham Young University - Hawaii, Laie, HI, in 1989, and the M.S. degree in Telecommunications and the Ph.D. degree in Information Sciences from the University of Pittsburgh, Pittsburgh, PA, in 1996 and 2006, respectively. Currently, he is an assistant professor of Telecommunications Systems Management at Murray State University, Murray, KY. His teaching and research areas include computer and network security, US telecom policies, network protocol analysis, network architecture design, QoS enabled networks, peer-to-peer networks, and video distribution networks. Dr. Okuda may be reached at masaru.okuda@murraystate.edu

MATRIX IMPACT ON THE RESIDUAL RESISTANCE FACTOR ESTIMATION OF POLYMER SOLUTIONS IN DUAL-POROSITY SYSTEMS: AN ANALYTICAL AND EXPERIMENTAL STUDY

Meysam Nourani, IPEX; Hamed Panahi, IPEX; Narges Jafari Esfad, Sharif University of Technology; Alireza Mohebi, IPEX; Ahmad Ramazani, Sharif University of Technology; Mohammad Reza Khaledi, IPEX

Abstract

The main focus of this study was The analytical and experimental determination of the residual resistance factor in a fractured medium taking into consideration both the matrix and fracture contributions to fluid flow. In a previous study, the authors showed that the residual resistance factor in a fractured medium was dependent on a dimensionless parameter called coil overlap, which is a function of both intrinsic viscosity and polymer concentration, and Power Law Equation parameter of a polymer [1]. However, as the experiments were conducted exclusively on glass micromodels, the matrix role was neglected. Therefore, in the analytical models developed for this study, for simplicity and practicality, it was assumed that the matrix flow contribution was negligible. The analytical solution of this theoretical model and the associated core-flood experiments showed that taking into account the role played by the matrix could significantly reduce error and improve the agreement between analytical and experimental results. Similar to the previous study, the parameters considered in this study were the polymer concentration, power law constitutive equation parameter, salinity, sulfonation content of the polymer, temperature, and molecular weight of the water-soluble polymers, which are used in polymer flooding for enhanced oil recovery.

Introduction

It is now considered common knowledge that substantial amounts of the remaining oil reserves are of types with very high viscosity and a range of gravities. However, the research community has a long history of working with crudes with high viscosity and relatively low API gravities [2], mostly in the clastic sedimentary rocks. Various EOR methods have been proposed for recovery of these crudes, among them thermal methods, microbial injections and chemical injections. Chemical injection methods comprise injection of surfactants, biosurfactants and polymer injection. What is lacking in this process is a study covering the recovery of crudes, which are not necessarily classified as heavy by polymer injection, from fractured carbonate rocks considering both the roles being played by fracture and matrix [3].

Various types of polymers have been utilized in the oil and gas industry. Partially hydrolyzed polyacrylamides (HPAM) and xanthan polysaccharides have been the leading polymers used in enhanced oil recovery (EOR). These two types are considered as the most cost effective types of polymers. Associative polymers have been investigated as a possible substitute for HPAM polymers in EOR applications. For hydrophobic associative polymers, incorporation of a small fraction of hydrophobic monomer into an HPAM polymer enhances intermolecular connections, thereby enhancing viscosities and resistance factors. At moderate concentrations, these polymers can provide considerably higher viscosities than polymers with equivalent molecular weights without hydrophobic groups [4].

Injection of a dilute solution of a water-soluble polymer, in this case partially hydrolyzed polyacrylamide (HPAM), to increase the viscosity of the injected water can increase the amount of oil recovered in some formations [5]. This polymer acts in three ways: 1) by impacting the fractional flow, 2) by reducing water mobility, and 3) by diverting the injected water to the non-invaded zones [2]. The ratio of the mobility of water to the mobility of a polymer solution under the same conditions is defined as the resistance factor (RF) [3] as follows:

$$RF = \frac{K_w}{K_p}\frac{\mu_p}{\mu_w} \tag{1}$$

whereas K_w and K_p are water- and polymer-relative permeabilities and μ_w and μ_p are water and polymer viscosities, respectively. Residual resistance factor (RRF) is another useful parameter, which is defined as the ratio of the initial water mobility to the water solution mobility after polymer flooding [3]. This definition, together with the definition of mobility ratio, which is mobility of water as the displacing fluid to mobility of oil as the phase being displaced, clearly explains the added value of polymers.

Enhancing the already established correlation between RRF and polymer concentration, water salinity, molecular weight of the polymer and temperature as the main objective of this study will be discussed in detail in proceeding sections. Eventually, residual resistance factor as a quantified parameter implicitly representing several characteristics of the polymer can serve in high-level studies for screening and identifying a fit-for-purpose EOR method.

Materials and Methods

Theory

In a previous study by Ramazani et al. [1], the following equation was developed correlating RRF, intrinsic viscosity and polymer concentration:

$$RRF = \exp\left([\eta]c - \ln\frac{4n_c}{3n_c+1} \right) \tag{2}$$

where n_c is a the power law constitutive equation parameter for the polymer concentration, c , and the intrinsic viscosity multiplied by polymer concentration is a dimensionless number called coil overlap parameter. Merely taking into account the significant role of the fracture as a flow conduit, the impact of matrix has been neglected in Equation (2), which potentially introduces a marked error in calculating the resistance factor. The attempt was made to factor the impact of flow into the matrix in the previous equation. To this end, in a model defined according to the conceptual model shown in Figure 1, the number of capillary tubes may be calculated with the following correlation for a core plug of diameter D with a longitudinal fracture of length L and an average opening of d .

$$n_f = \frac{L}{d} \tag{3}$$

The assumption to treat a 2-dimensional (actually, 3-dimensional) fracture with a series of one-dimensional tubes is true only when the fracture thickness in the 3rd dimension does not exceed a certain amount. This somehow falls into the category of micro fractures. The total flow rate of the core is equal to the flow rates of matrix and fracture. So, Equation (3) can be expanded for total flow rate (Q_t), matrix flow rate (Q_m) and fracture flow rate (Q_f):

$$Q_t = n_f Q_f + Q_m \tag{4}$$

Replacing the flow rates from the Darcy equation:

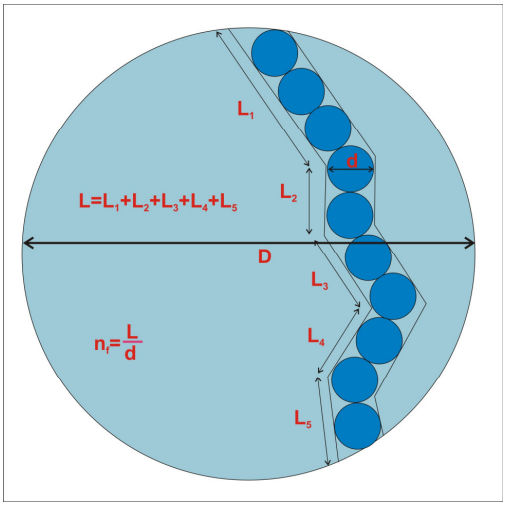


Figure 1. Schematic of the Fracture and Matrix Network in the Conceptual Model

$$-\frac{\pi}{4}D^2\frac{k_t}{\mu}\frac{\partial p}{\partial l} = -n_f\frac{\pi}{4}d^2\frac{k_f}{\mu}\frac{\partial p}{\partial l} - \frac{\pi}{4}(D^2 - n_f d^2)\frac{k_m}{\mu}\frac{\partial p}{\partial l} \tag{5}$$

in which k_t , k_f and k_m are total, fracture and matrix permeabilities. For flow in a dual-porosity medium, Equation (5) holds when the system is not far from equilibrium. The system reaches equilibrium as soon as the injection ceases. Equation (5) can then be simplified to:

$$k_t = n_f\alpha k_f + (1 - n_f\alpha)k_m \tag{6}$$

in which α is $(\frac{d}{D})^2$.

Flow rate of a polymer solution with viscosity μ and the power law constitutive equation parameter n in a capillary tube with radius r and length l are correlated to the pressure drop according to Equation (7).

$$Q = -\frac{n}{2(3n+1)}\frac{\pi R^4}{\mu}\frac{\partial p}{\partial l} \tag{7}$$

Assuming that the fracture consists of n_f capillary tubes, and comparing Equation (7) with Darcy's law, Equation (8) holds for a capillary tube containing a polymer solution.

$$k_{fp} = \frac{nd^2}{8(3n+1)} \tag{8}$$

The assumption to treat a 2-dimensional (actually, 3-dimensional) fracture with a series of one-dimensional tubes is true only when the fracture thickness does not exceed a certain amount. As water is acting as a Newtonian fluid, the power law constitutive equation parameter is equal to one. So, for the case of a fracture containing water. Equation (8) simplifies to:

$$k_{fw} = \frac{d^2}{32} \tag{9}$$

Writing Equation (6) for water and polymer solutions separately and replacing the permeability for water and polymer solutions in a fracture from Equations (8) and (9), one has:

$$k_{tw} = \frac{n_f \alpha d^2}{32} + (1 - n_f \alpha) k_{mw} \tag{10}$$

$$k_{tp} = \frac{n_f \alpha n d^2}{8(3n+1)} + (1 - n_f \alpha) k_{mp} \tag{11}$$

Equating Equation (10) to Equation (11) while multiplying by the ratio of polymer solution viscosity to water viscosity, which is called relative viscosity,, η_{rel} gives the following correlation for the residual resistance factor:

$$RRF = \frac{\frac{n_f \alpha d^2}{32} + (1 - n_f \alpha) k_{mw}}{\frac{n_f \alpha n d^2}{8(3n+1)} + (1 - n_f \alpha) k_{mp}} \eta_{rel} \tag{12}$$

Rewriting Equation (12) for relative viscosity yields

$$\eta_{rel} = \frac{\frac{n_f \alpha n d^2}{8(3n+1)} + (1 - n_f \alpha) k_{mp}}{\frac{n_f \alpha d^2}{32} + (1 - n_f \alpha) k_{mw}} RRF \tag{13}$$

Taking the derivative of the natural logarithm of Equation (13), while taking into account the definition of intrinsic viscosity, leads us to

$$\left.\frac{\partial \ln RRF}{\partial c}\right|_{c \to 0} = [\eta] - \left.\frac{\partial}{\partial c} \ln \frac{\frac{n_f \alpha n d^2}{8(3n+1)} + (1 - n_f \alpha) k_{mp}}{\frac{n_f \alpha d^2}{32} + (1 - n_f \alpha) k_{mw}}\right|_{c \to 0} \tag{14}$$

Integrating Equation (14) with respect to polymer concentration, the residual resistance factor as an exponential function becomes

$$RRF = \exp\left([\eta] c - \ln \frac{\frac{n_f \alpha n_c d^2}{8(3n_c+1)} + (1 - n_f \alpha) k_{mp}}{\frac{n_f \alpha d^2}{32} + (1 - n_f \alpha) k_{mw}} \right) \tag{15}$$

If matrix does not play any role in the flow, it can be assumed that the number of capillary tubes, n_f , and the coefficient, α , are equal to one. Then, Equation (15) becomes

$$RRF = \exp\left([\eta] c - \ln \frac{4 n_c}{3 n_c + 1} \right) \tag{16}$$

And, if the ratio of fracture opening to matrix/plug diameter, α, is negligible, then

$$RRF = \exp\left([\eta] c - \ln \frac{k_{mp}}{k_{mw}} \right) \tag{17}$$

Experimental Polymers

Similar to the previous study by the authors for the experiments on the micro-models, six well-characterized polyacrylamides with different molecular weights and sulfination levels were used here. The specifications of these polymers are listed in Table 1.

Table 1. Specifications of the Polymers used in the Core-Flood Experiments

No	Name	Mw.10^{-6} (Dalton)	Sulfonation (%)	Hydrolization(%)
1	PAAM 30	30	0	0
2	HPAAM 20	20	0	0-25
3	HPAAM 8	8	0	0-25
4	PAAMS 832	8	32	0
5	PAAMS 825	8	25	0
6	PAAMS 65	6	5	0

The desired salinity of the solutions was adjusted using NaCl with a purity of at least 99.5%. The water used was double distilled with an all-glass apparatus. Polymer solutions with the required concentrations were prepared by

slowly dissolving the polymer in the distilled water. To avoid any aging effect, the solutions were gently stirred on a mechanical shaker for about 24 hours.

Crude Oil

The surface oil sample of the S-Field which was taken at the well head and used for polymer flooding in micro-models was also used for core-flood experiments. This is relatively light crude with an API gravity of 31.2.

Apparatus

Rotational viscometer

The variation of shear stress of the polymer solution with polymer concentration and shear rate were evaluated with the Physica MCR 301 rheometer.

Capillary Viscometer

Dilute solution viscosity measurements were made using an Ostwald-type viscometer, size 250, manufactured by Petrotest. The viscometer was mounted vertically in a thermostatic bath at appropriate temperatures controlled with a precision of 0.1ºC. The flow times were recorded with a stopwatch capable of registering the time with a precision of 0.1s. Each test was consecutively repeated at least three times and the average of three flow times was recorded.

Sample Acquisitions and Cores

A suitable core plug was used for these experiments. This core plug was from a wellbore of an offshore field with the reservoir conditions listed in Table 2.

Table 2. Plug Descriptions

No.	Depth (m)	KL (mD)	Grain Density (g/cc)	Porosity (%)	Plug Description
6	2930.7	127.969	2.717	3.003	Fractured, Limestone

The offshore fields are the best candidates for polymer injection due to the fact that the essential ingredient of polymer solutions is water. The core-flood experiments were conducted at the reservoir's temperature and pressure. The plug was first cleaned and prepared for primary saturation with a brine of the reservoir's salinity. Then, it was flooded with the oil sampled from the same reservoir to simulate the oil migration process after which it was flooded with brine of a salinity close to sea water. This step represents the water flooding process prior to polymer injection as the field of interest was water flooded. This was followed by a slug of 0.3 pore volume and 700ppm polymer solution and the subsequent brine injection until almost no more oil could be attained. During all flow experiments, the RRF values were recorded by measuring the initial water mobility to the water solution mobility after polymer flooding.

Results and Discussions

The power law constitutive equation parameter was determined using the slope of plotting the logarithm of shear stress as a function of logarithm of shear rate. The calculated power law constitutive equation parameters, for all polymers at 700ppm polymer concentration but at the temperature and salt content of the reservoir, are presented in Table 3.

Table 3. Core-Floods Experimental Conditions

Overburden Pressure (psi)	Pore Pressure (psi)
9500	3500
Temperature (C)	Res. Salinity (ppm)
109	154400
Sea water Salinity	Average depth (m)
41542	28848
Oil Density (g/cc)	Water Injection rate (cc)
0.897101	1.2

To calculate the resistance factor at temperatures and salinities of the reservoir, the intrinsic viscosity had to be estimated under those conditions. As it was not possible to measure intrinsic viscosity at 109ºC, hydrodynamic diameter of polymers at temperatures and salinities of the reservoir had to be estimated from Equation (18). Assuming that the polymers are spherical, and with the help of Equation (19), the intrinsic viscosity for polymer solutions could be calculated for reservoir conditions of

$$\begin{aligned} LnD = (&-2.39*10^{-2}*C_{Salt} \\ &-2.6*10^{-4}*T+5.06*10^{-3}C_{Sulf.} \\ &+3.40*10^{-2}*M_w-15.0939) \end{aligned} \tag{18}$$

where M_w is molecular weight, T is temperature, C_{Salt} is salt concentration, and $C_{Sulf.}$ is sulfonation content of the polymer, such that

$$[\eta] = \frac{\pi . N . D^3}{6 M_w} \quad (19)$$

where N is the Avogadro constant and D is the hydraulic diameter of polymer molecules in meters. Using glass micro-models and the simplified analytical solution, it was previously shown that the recovery factor is logarithmically proportional to the resistance factor [1]. The calculated intrinsic viscosities can be found in Table 4.

Table 4. Estimated Polymer Viscosities

Polymer	Estimated Diameter of Polymer (m)	Estimated Intrinsic Viscosity (dL/g)
PAAM 30	4.6546E-07	10.59
HPAAM 20	3.31301E-07	5.73
HPAAM 8	2.20308E-07	4.21
PAAMS 832	2.59031E-07	6.85
PAAMS 825	2.50016E-07	6.16
PAAMS 65	2.59031E-07	9.13

Thin sections of the core plugs were used for the estimation of α defined for Equation (6). Figure 2 illustrates a fracture opening in plug 6. After analyzing the thin sections, an average opening of 53 micrometers was assumed for plug 6.

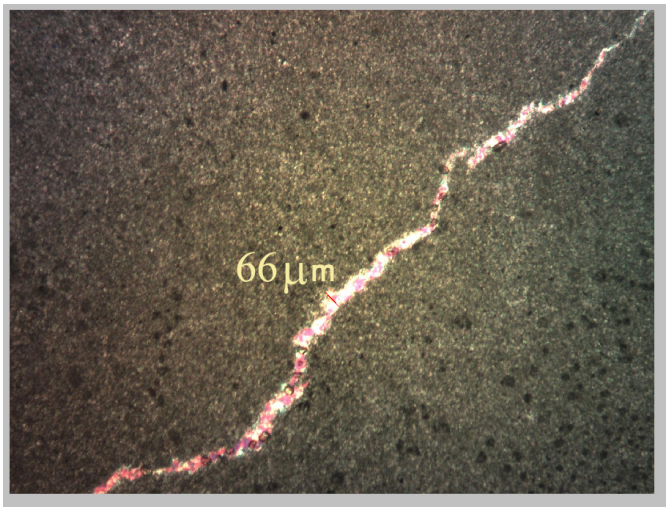


Figure 2. Fracture Opening in Plug 6

The analysis further showed a fracture width of 0.00762m for the plug 6. So, n_f was calculated to be 144 for this case. From Equations (10) and (11), relative matrix permeabilities can be calculated for the water and polymer solutions, the result of which is shown in Tables 5 and 6. The last two columns show the calculation errors for these cases. Clearly, the error is reduced dramatically when both the fracture and matrix contributions are taken into account.

Table 5. Estimated Polymer Viscosities (1)

Polymer	Cal. RRF (Fr.+Mat.)	Cal. RRF)Fr.)	Measured RRF	Er (%) (Fr.+Mat)	Er (%) (Fr.)
HPAAM 20	2.21	1.54	2.46	10.37	37.43
HPAAM 8	1.42	1.36	1.44	1.06	5.26
PAAMS 832	2.08	1.66	2.21	6.02	24.87
PAAMS 825	1.80	1.58	1.86	3.35	15.19
PAAMS 65	1.26	1.44	1.22	-2.99	-17.67
PAAM 30	4.25	2.20	5.48	22.40	59.88

Table 6. Estimated Polymer Viscosities (2)

Polymer	nc	Ktw (mD)	Ktp (mD)	Kmw (mD)	Kmp (mD)
HPAAM 20	0.89	127.97	86.23	103.25	62.59
HPAAM 8	0.94	127.97	120.41	103.25	96.38
PAAMS 832	0.90	127.97	99.17	103.25	75.44
PAAMS 825	0.91	127.97	109.36	103.25	85.53
PAAMS 65	0.94	127.97	123.41	103.25	119.43
PAAM 30	0.84	127.97	62.88	103.25	39.56

Conclusions

In accordance with the theoretical and experimental studies, an analytical relationship—to calculate RRF in fractured medium considering the impact of both matrix and fracture—was developed, which correlates RRF with the coil overlap parameter and water and polymer solution permeabilities in matrix for the case of small fractures relative to matrix size (see Equation (17)). The RRF estimation error dropped significantly, on the order of 3 to 5 fold; refer to Tables (5) and (6).

The experimental results presented here are in agreement with results of polymer flooding in micro-model experiments, confirming that ultimate oil recovery in the polymer flooding process increases as RRF increases. Flow in a dual-porosity system can be routinely treated by the continuum methodology such as the double-permeability or dual-porosity approaches or the discrete fracture network (DFN)

models with permeable walls (porous matrix), but the approach presented here simplifies the technique significantly, while not sacrificing results.

Nomenclature

RRF = residual resistance factor
RF = resistance factor
HPAM & PHPA = Partially Hydrolyzed Polyacrylamide
DFN = discrete fracture network
K_w = relative permeability of water
K_p = relative permeability of polymer solution
μ_w = water viscosity
μ_p = polymer solution viscosity
n_c = power law constitutive equation parameter
D = core plug diameter, hydraulic diameter of polymer molecule in meters
L = fracture length
d = fracture opening
Q_m = matrix flow rate
Q_f = fracture flow rate
Q_t = total low rate
k_t = total permeability
k_f = fracture permeability
k_m = matrix permeability
η_{rel} = relative viscosity
M_w = molecular weight
T = temperature
C_{Salt} = salt concentration
$C_{Sulf.}$ = sulfonation content of polymer
N = Avogadro constant

References

[1] Ramazani, R., Nourani, M., Emadi, M. & Esfad, N. (2010). Analytical and Experimental Study to Predict the Residual Resistance Factor on Polymer Flooding Process in Fractured Medium. *Transport in Porous Media*. Springer Journal # 9594.

[2] Riley, B. N. & Peter, H. D. (1987). Polymer flooding review. *Journal of Petroleum Technology*, 39(12), 1503–1597.

[3] Jennings, R. R., Rogers, J. H. & West, T. J. (1971). Factors influencing mobility control by polymer solution. *Journal of Petroleum Technology*, 23(3), 391–401.

[4] Seright, R. S., Fan, T., Wavrik, K., Wan, H., Gaillard, N. & Favéro, C. (2011). Rheology of a New Sulfonic Associative Polymer in Porous Media, *SPE International Symposium on Oilfield Chemistry*, 11-13 April 2011, The Woodlands, Texas, USA

[5] Dominguez, J. G. & Willhite, G. P. (1977). Retention and Flow Characteristics of Polymer Solutions in Porous Media. *SPE Journal*, 17(2) 111-121.

Acknowledgments

The authors would like to acknowledge the influence that Mr. A.H. Sharifi, MAPSA managing director, has had on the inspiration and preparation of this paper.

Biographies

M. NOURANI , the corresponding author, has a PhD in petroleum engineering from Sharif University of Technology, Iran and studied chemical and petroleum engineering at Tehran University and Sharif University. He has been in the oil industry since 1998.Also he is the director of MAPSA's laboratories. Dr Nourani may be reached at Nourani@mapsaeng.com

H. PANAHI has been involved in wide range of proprietary and non-proprietary research projects from polymer injection to modeling of shale behaviors in research centers in Iran, Europe and the US in more than a decade. He may be reached at hpan@statoil.com

INFORMATION TECHNOLOGY & HEALTH: A NEW ARENA IN THE HOSPITAL INDUSTRY

Abdulrahman Yarali, Murray State University; Daniel Claiborne, Murray State University; Solomon Antony, Murray State University

Abstract

Applications of information technology can be found in all aspects of a modern economy and the healthcare field is no exception. Health Information Technology (HIT) consists of an enormously diverse set of technologies for transmitting and managing health information for use by the consumers, insurers and other groups with an interest in health and healthcare. With healthcare costs and quality assurance taking central roles in the healthcare arena, increasing attention is being directed towards the potential of health IT to lower healthcare spending and improve efficiency, quality and safety of medical care delivery. One of the primary motivators for adopting health IT applications is the belief that they improve the quality of patient care irrespective of where these services are delivered. The movement of telemedicine from fixed platform to mobility and wireless infrastructure can have a significant impact on the future of healthcare. This paper presents a summary of Healthcare Information Technologies (HIT), applications, advantages, challenges and issues.

Introduction

With only half of all Americans receiving care that meets clinical quality standards, healthcare quality needs to be improved by accelerating smart investments in Health Information Technology (HIT). Such investments will ensure that providers and their patients receive better and timely access to key healthcare data. HIT provides a framework to describe the comprehensive management of health information and its secure exchange between consumers, providers, government and quality entities (a public or a private entity that is qualified to use claims data to evaluate the performance of providers and suppliers on measures of quality, efficiency, effectiveness and resource use, and that meets the eligibility requirements enumerated in the proposed rule) and is in general increasingly viewed as the most promising tool for improving the overall safety and efficiency of the healthcare delivery system. HIT can help prevent medical errors, improve care coordination, increase access to providers in rural areas and enhance the overall value of care. HIT includes a variety of integrated data sources including patient Electronic Medical Records (EMR), Decision Support Systems, and Computerized Physician Order Entry (CPOE) for medications. Creating and maintaining such systems is complex; however, the benefits can include dramatic increases in efficiency savings, greatly increased safety, and health benefits.

Information systems used in the healthcare industry can be studied using the popular IT infrastructure framework [1]. Health information systems mostly support, track and evaluate the delivery of healthcare. The basic system upon which everything is built is the EMR. Like today's paper-based medical record, EMR includes the patient's history, diagnoses, tests that were ordered and test results, prescriptions, physician's comments, and, in the most complete form, x-rays and other medical images. However, unlike today's paper records, EMRs can be easily shared and analyzed [2].

As information technologies continue to evolve, the skills that are necessary to employ are equally turning out to be more and more sophisticated. The result of this is that as availability of technology does continue to grow, the risk of misinformation, misused information and missed information is also expected to rise, potentially leading to dissatisfied users and poor quality of healthcare.

In subsequent sections of this paper, the authors present an overview of some of the technologies used in HIT, details on some of the consequences of application of the technologies, challenges in HIT adoption, ethical implications of using HIT, the status of the industry and guidelines for future research.

Health Information and Advanced Technologies

Applications of information technology can be found in all aspects of a modern economy. Information systems are comprised of software, hardware, communication and collaboration networks, data facilities and human resources [1]. The HIT arena can be analyzed using this framework. The software used in the healthcare industry can be classified based on the functional the software serves. Finance-related software includes packages for medical billing and insurance management. Patient-care-related software packages include case management, patient scheduling,, information

management related and patient care administration. Use of integrated clinical information systems that facilitate communication between various systems—such as physician order entry, pharmacy and laboratory information systems, clinical decision support systems and clinical drug disposal systems—leads to a decrease in medication errors, and improved medication administration safety [3]. Implementation of clinical information systems has resulted in nurses spending more time on nursing care than before [4].

The use of HIT has also empowered patients, as they are better able to gain access to health information without necessarily having to depend on a physician's advice or having to visit hospitals and clinics. Patients also have the advantage of being able to select what they are going to hear and read about healthcare [5]. Although health information remains easily accessible and readily available to almost everyone, cost, difficulty of implementation and other barriers impede adoption of such systems, and studies have documented low rates of technology acquisition and implementation in emergency departments (ED) and other settings. Studies of HIT adoption in physicians' offices and hospitals throughout the country also revealed a slow uptake by 2008 [6] .These studies found that only 7.6% of U.S. hospitals have a basic electronic records system, with only 1.5% having a comprehensive one [7]. It was also found that hospitals that have academic affiliations and that have high IT operating budgets and staff tend to have more automated clinical information systems [8]. Use of integrated clinical information systems does not guarantee error-free operations; however, downtime due to hardware errors, software errors, patches and user errors can be a serious cause for concern and can lead to medication errors [9].

In our modern world, the force behind healthcare is mainly being driven by the great need for enhancing access to the use of HIT, irrespective of where these services are to be delivered. It has been noted by several researchers that modern technological innovations, especially in IT and telecommunication systems, have increasingly influenced our standards of healthcare, mainly by allowing both the providers and patients to be in a position where they are better informed. The adaptation and acceptance of IT solutions among the users of clinical information systems tend to vary by user groups. It was found that Australian nurses and midwives expressed predominantly negative experiences with computerized patient information systems [9]. General practitioners, on the other hand, tend to have high interest in the use of systems when they are first implemented, though their interest wanes over time [10]. It was also found that the physicians in Norway preferred the paper-based system during the patient discharge process, using EMR systems only for background information and verification [11].

Following is a list of the most promising health technologies which have been considered by the experts[12]:

- Instant Medical Data Collection and Knowledge Dissemination Technologies and Standards
- Decision Making and Support Technology (personal and point of care)
- Individualized Diagnosis and Treatment (e.g., real-time protein synthesis, real-time genetic testing)
- Health Systems Methodologies
- High Tech Intervention (e.g., robotic surgery, sensors, tele-consultations)
- Information Access and Feedback Technologies
- New Technology Evaluation Methodologies

Telehealth/telemedicine refers to the delivery of health-related information and services through telecommunication technologies, which may include healthcare education. The aim is to provide expert-based medical care anywhere healthcare is needed. Telemedicine applications, including those based on wireless technologies, span the areas of emergency healthcare, tele-cardiology, tele-radiology, tele-pathology, tele-dermatology, tele-ophthalmology, tele-oncology and tele-psychiatry [13].

With healthcare costs soaring, policymakers are looking for ways to streamline the administration and cost of healthcare services. A key platform for achieving this objective is broadband. Indeed, broadband is driving innovation and spurring cost-savings in the healthcare sector by providing a robust, interactive medium that enables a variety of telemedicine tools and services by facilitating anytime-anywhere computing. The impact of these tools and services is evidenced in the following examples:

- Enabling the use of efficient HIT. Broadband enables the widespread use of electronic health records, which could streamline the administration of healthcare and lead to annual cost savings of approximately $80 billion [13], [14]. In addition, coordination between various players in the delivery of care can be enhanced by using RFID technologies [15], [16].
- Enhancing the quality of care. The use of broadband-enabled telemedicine and HIT tools can reduce costly medical errors via the implementation of solutions like e-prescribing, which can enhance physician accuracy [17].
- Extending the geographic reach of healthcare to rural areas. The difference in the quality of healthcare available in rural and urban areas is significant. However, broadband is being used to enable tele-consultations, tele-radiology and remote monitoring, all of which help to make up for a dearth of physicians who practice in rural areas.

- Facilitating in-home care. Broadband-enabled in-home monitoring systems are increasingly popular among seniors, people with disabilities and others. These types of services could enable enormous cost-savings. According to one study, "a full embrace of remote monitoring alone could reduce healthcare expenditures by a net of $197 billion (in constant 2008 dollars) over the next 25 years with the adoption of policies that reduce barriers and accelerate the use of remote monitoring technologies" [18].
- Reducing unnecessary treatments and costly patient transfers. A pilot initiative sponsored by the U.S. Veterans Association found that in-home chronic-disease management tools (e.g., tele-consultations, remote diabetes monitoring) resulted in 40% fewer emergency room visits and a 63% reduction in hospital admissions [19].
- More cost-effective healthcare for seniors and people with disabilities. According to one estimate, broadband -enabled health and medical services can save some $927 billion in healthcare costs for seniors and people with disabilities [20].

The movement of telemedicine from fixed platform to mobility and wireless infrastructure can have a significant impact on the future of healthcare. Cutting-edge innovations in mobile health technology have the potential to help overcome the gaps between health and care, enabling patients to take a more active role in their healthcare and connecting physicians to vital, real-time information that supports improved treatment and preventive care. The fundamental advantage of this small wedge would be

- to empower patients by putting more data in their hands, and enabling them to make more data/ evidence-driven decisions, and
- to enable care providers (doctors, nurses, insurance companies and hospitals) to track a patient's progress in-between visits, thereby providing much more granular measurements on how a patient might be responding (over time, to prescriptions, etc).

Mobile technology can enable real-time monitoring in a way that was prohibitively expensive just a few years ago. Real-time communication can enable care providers to recognize and respond to health issues rapidly, and to provide more data to drive medical advice and recommendations, and help all parties make decisions.

Tele-health remains one of the means through which patients in rural locations can gain access to healthcare information, especially when the most needed care and services happen to be a good distance away. Though telehealth does offer promise, accessibility and location remain problematic to some patients who cannot travel even short distances. Though there is a lot of data supporting the use of telehealth, there are several physicians who still avoid using electronic triage systems as they believe that the experience, wisdom, knowledge and skills of a physician remain the gold standards for providing appropriate healthcare. This can also be coupled with the fact that all physicians have a clear professional obligation to make the best use of their knowledge in offering optimal care to all patients [21].

Electronic medical records refers to the use of patient records that are computerized. The EMR's structure as a store of electronic information, capable of being searched, categorized and analyzed, makes it superior to the traditional paper chart for informing those in charge of the care process. Nevertheless, proceeding from its historical basis as the digital version of a patient's chart, the EMR is a provider-focused view of the patient's health history. It comprises health-related information that is created by clinicians or that results from clinician orders and activity on behalf of a patient, such as diagnostic tests or prescriptions for medications. A main objective of an EMR is to improve the ability of a clinician to document observations and findings and to provide more informed treatment of persons in his or her care. The EMR contains demographic information and clinical data on the individual, including information about medications, the patient's medical history and the doctor's clinical notes. Because of the lack of interoperability, an EMR is limited to one healthcare organization. This does not mean a single physical location; under some circumstances, information can be shared among multiple facilities and still be within one EMR

Electronic clinical support systems are a type of knowledge-based technology used to support most of the clinical decision-making processes starting from the point of diagnosis and continuing to investigations to be carried out and treatment offered and recovery options recommended.

Online healthcare resources refers to all web-based resources that provide information to both healthcare consumers and providers. Some of the information provided here may include—though not be limited to—product availability, dental and medical services, hospitals, providers, alternative healthcare options, publications, employment and mental health. There is an increasing wealth of information that is available to most people, even those in remote areas. In most cases, individuals who go online searching for information are mainly seeking advice, making them vulnerable to misinformation [21].

In our modern free society, nearly anyone is free to publish opinions and information on the Web; judgment of the

site's reliability is left to the individual user. The use of online research, therefore, turns out to be a challenging enterprise for anyone seeking healthcare. Since health information is increasingly being marketed, it is to be expected that there will be a lot of commercial influence on what information is conveyed. Information available on the Internet may also express unilateral and biased opinions of a certain organization or group. Information of this nature can, thus, be potentially erroneous, misleading or misinterpreted, and can very easily cause inappropriate or even harmful decisions to be made by the support team [6].

Applications

HIT is applied in the provision of healthcare services to improve efficiency, quality, cost-effectiveness as well as the safety of all medical care procedures in every nation's healthcare sector [6]. Appreciating these benefits remains extremely important, especially where reports clearly show that there has been a steady annual increase in the cost of health intervention within the last five years. Equally noted by most of these reports is that more than 50% of all cost is wasted on misdiagnoses or on inefficient processes [5]. The outcry of legislators and other organizational leaders has been to emphasize the importance of using computer-based information technology to enhance accessibility of critical information, to minimize human error and to automate inefficient and labor-intensive processes.

The most appropriate use for HIT is to help medical practitioners minimize medical errors [5]. Technology-based strategies in most of the cases, when used appropriately, have proven to be effective in eliminating human error in several industries such as aviation and banking. Equally, the use of clinical HIT systems has resulted in substantial increases in medical safety and quality by making use of the most relevant automated system in decision making. Technology also comes in handy when practitioners are seeking knowledge acquisition, thus reducing errors of omission.

In the case of an environment where ambulatory healthcare is needed, use of HIT does offer various benefits. First of all, it improves the financial aspect and efficiency of the entire practice. For years now, several offices making use of computerized financial and scheduling systems were helped in streamlining various office processes. The computerized system helps in tracking the productivity of the entire practice as well as in automating the reimbursement processes. Secondly, use of these ambulatory EHR systems offers a good opportunity to improve and monitor clinical quality delivered mainly by improving access to information and also in helping to reduce duplicative documentation. The use of EHR comes with the advantage of allowing system connectivity in different departments of a hospital and the exchange of information among different providers from the same organization, different organizations and even nationally [5].

Though there are several advantages associated with the use of technology, there are still several medical providers and organizations that have not been willing to fully adopt HIT. A recent survey that sought to find out the use of computerized physician order entry (CPOE) found that only 9.6% of all of the hospitals in the developed world have completely incorporated CPOE in their system; and from these, only half of them demanded that CPOE had to be used. In the case of ambulatory settings, some recent estimates placed the use of electronic health records at only 6-15% of all office-based physicians [6]. The huge potential advantages of widely adopting HIT in the healthcare system does necessitate that any scientific evidence supporting benefits of HIT-related costs be examined. Also to be evaluated in this kind of a case are the potential barriers that exists when an organization is trying to implement various types of HIT systems in its effort to provide a better healthcare environment [21].

Several reviews have pointed to the huge potential that HIT has for dramatically transforming the delivery of healthcare services by making them safer, more efficient and more effective. To be sure, however, the evidence of empirical research supporting these HIT benefits is still limited, thus calling for more research in this area. Irrespective of the particular context, the impact of implementing HIT on quality and cost has been shown not to be consistent across all institutions. This is because the specific context upon which HIT is often implemented is affected by factors such as setting, patient population and the clinic. A more widespread implementation of HIT has, in most cases, been limited by a lack of knowledge as to what methods for implementation and types of HIT are best suited for particular organizations to give the best results, especially for small hospitals and small practices. To be able to derive maximum benefit from HIT, reports of HIT implementations and developments ought to be improved. Greater attention should be paid to how the descriptions of its intervention fit the organizational or economic environment.

Since HIT is turning out to be extremely famous and nearly anyone can write something on it, there is still a need to come up with standards for any information that is delivered, as is the case with other standards that demand clinical trials for therapeutics before they can be released to the general public. While making use of existing evidence that has been published, it remains difficult to come up with conclusions as to which HIT functionalities could be best suited

for achieving specific health benefits. The assessment of the probable costs likely will be even more difficult. It should be noted that the existing evidence happens not to be sufficient when one is trying to answer the questions "who benefits from what?" and "who pays for what?" Researchers in this area can develop appropriate models that can estimate the potential benefits and costs of implementing an HIT system across particular healthcare settings [5]. Several smaller high-quality studies have shown that the implementation of an ambulatory EHR system clearly improves the quality of patient care.

Benefits and Challenges of HIT

Evidence of the Benefits of HIT

Since the most important thing in healthcare is patient information, all of this information is stored in the form of EMR. Using such records, hospitals can eliminate searches for medical charts or papers, give the appropriate drugs, identify harmful drugs and prescribe the right ones. The Consumer Empowerment and Access to Clinical Information via Networks Interoperability Specification defines specific standards needed for assisting patients in making decisions regarding care and healthy lifestyles. The Emergency Responder Electronic Health Record Interoperability Specification defines specific standards required for tracking and providing on-site emergency-care professionals, medical examiner/fatality managers and public health practitioners with needed information regarding care, treatment or investigation of emergency-incident victims.

The Medication Management Interoperability Specification defines specific standards for facilitating access to necessary medication and allergy information for groups such as consumers, clinicians, pharmacists, health insurance agencies, inpatient and ambulatory care offices, etc. The Quality Interoperability Specification defines specific standards needed for benefiting providers by providing a collection of data for inpatient and ambulatory care and for benefiting clinicians by providing real-time or near-real-time feedback regarding quality indicators for specific patients. The integration of electronic records that can communicate with each other, governance and oversight organizations, and health information exchange processes, will establish a larger and fully connected infrastructure to support all aspects of health and care.

Various data and technological standards currently in use are proving integral in the development of "interoperable" health information systems capable of effectively sharing health data included in electronic health records and electronic prescribing. One main type of standard lays out a common set of medical terminology for a particular area of healthcare, in order to help ensure that all information users understand one another. Another main type of standard spells out the uniform technical specifications that allow different computer systems to communicate accurately with one another. One popular standard in this category is known as Health Level Seven (HL7), a "messaging" standard that allows users to know who is sending and receiving the information and which patient the information describes.

The technical infrastructure that supports each of the public, private and domain-specific health-information exchanges fall into one of three categories: Federated, Centralized or Blended [22]. Under this approach to sharing medical data, each participating health entity, such as a doctor's office, hospital or lab, stores the data pertaining to its patients on its own separate computer system. These individual systems are then linked by a computer network that allows users to search for health records on each of the other systems using patient-indexing and record-locator software. Each participating health entity can maintain different computer programs at its own location as long as those programs can communicate with each other. An example of a hypothetical federated RHIO is shown in Figure 1. Here, HIE stands for Health Information Exchange network. It is the information technology structure that enables health data transfer. EHR stands for Electronic Health Record. Data is stored at each provider location, not in a central location. PHR stands for Personal Health Record, which enables individuals to access their health records. PI/RL stands for Patent Index and Record Locator software. These tools guide data requests through the network to the relevant information about the correct patient.

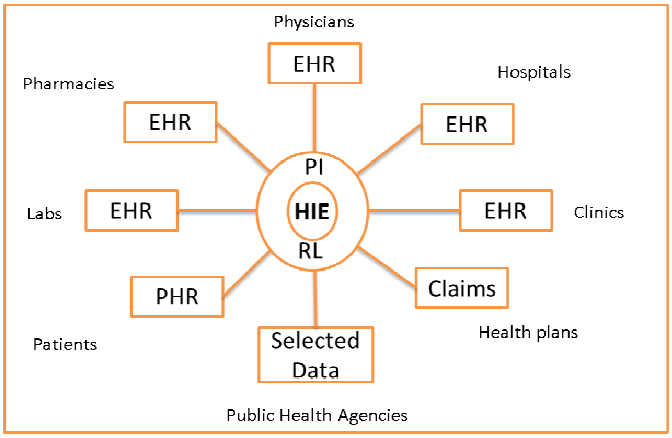


Figure 1. Regional Health Information Organization Federated System Example

Healthcare visionaries foresee a time when all types of health-related information exists electronically and can be

reliably and securely accessed by any number of authorized parties and entities to improve the health of an individual, a specific community or the U.S. population as a whole. The integration of electronic records that can communicate with each other, governance and oversight organizations, and health information exchange processes, will establish a larger and fully connected infrastructure to support all aspects of healthcare. While electronic records of health information are distinct entities now, it is clear that at some point in the future data within them could meld, and various portals or views to the information would be developed to support the needs of providers, individuals, researchers, public health officials and others engaged in health and wellness. This will bring additional benefits such as personalized healthcare, knowledge management and expectation of quality. The following is a graphic representation of how the components of the proposed infrastructure are integrated [23].

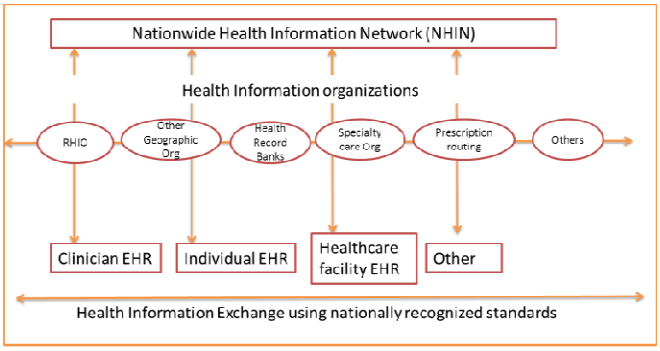


Figure 2. A Graphic Representation of how the Components of the Proposed Infrastructure Integrate [23]

Two studies, one by the RAND Corporation and one by the Center for Information Technology Leadership [24], [25], report estimates of the potential net benefits that could arise nationwide if all providers and hospitals adopted health information technology and used it appropriately. Both studies estimated annual net savings to the healthcare sector of about $90 billion, relative to total spending for healthcare of about $2 trillion per year. The studies, however, measured different sources of such savings. The RAND research focused primarily on savings that the use of health IT could generate by reducing costs in physicians' practices and hospitals, whereas the CITL study limited its scope to savings from achieving full interoperability of health IT, explicitly excluding potential improvements in efficiency within practices and hospitals.

Challenges

According to a 2011 survey, the deployment of Electronic Health Records systems appears to be a major item on the agenda of more than half of the hospitals surveyed. The hospitals apparently do not foresee potential problems in the complicated process of implementing a new system that complies with federal requirements [26].

Policy. According to experts, the US Government's policies for promotion of HIT adoption must be coordinated with a broader healthcare reform policy [27].
Government funding. Almost non-existent government funding for HIT has resulted in lack of HIT adoption in government health facilities and a lack of trained medical informatics professionals.
Computer literacy. Low computer literacy among government staff and, to a large extent, in the private provider community is a concern.
Infrastructure and coordination. Lack of supporting infrastructure and coordination between public and private sectors will take time to be resolved.
Legacy systems. Except for a very few privately owned large hospitals, most patient records are paper-based and very difficult to convert to electronic format. According to a recent survey in the U.S. ,more than 50% of the hospitals have moved away from legacy systems and into EMR systems. However, the adoption rate is much lower for smaller practices [28].
Standards. Some local HIT systems do not adhere to standards for information representation and exchange. This could be further complicated because of the use of multiple local languages by patients and some health workers.
Privacy. Patients are sensitive to disclosing their health information online because of privacy concerns and their personal dispositions [29]. Patient confidentiality can be ensured with incentives for compliance and disincentives for non-compliance as is the case in the U.S.
Cost. Costs include the initial fixed cost of the hardware, software and technical assistance necessary to install the systems.
Licensing fees. The expense of maintaining the system and the "opportunity cost"—time that the healthcare providers could be spending seeing patients, but instead must devote to learning how to use the new system and how to adjust their work practices accordingly.

HIT and Ethics

The current common use of HIT is changing the way medical providers take care of patients on a day-to-day basis. This has changed the efforts of medical practitioners to promote and support decision-making processes even in rural areas. Though technological interventions have been widely accepted in the modern set up, its use in remote settings has raised some questions about a conflict of ethics. It is the complex patient information and history, shortage of

service, increasing medical interventions options and treatments as well as high demand for medical services that are constantly challenging healthcare providers as they work to maintain appropriate health standards [5]. Even though the intentions of using these technologies are good, there should be extra scrutiny in all areas where it is applied so as to protect the welfare of the patients. When HIT is to be deployed in any setting, whether urban or rural, all healthcare providers ought to put patient welfare above any other considerations so as to protect his/her confidentiality, promote trust and ensure privacy in the entire healing relationship.

Trends and Future Directions

Ethically, this remains to be highly beneficial to the patient as long as all of the information that is obtained by the patient is appropriate, accurate and verifiable, and does not harm the patient psychologically in any way. For healthcare providers, use of modern electronic sources of information on a day-to-day basis remains unavoidable. It is now estimated that there are only a few healthcare interventions in the modern world that do not either indirectly or directly make use of HIT. In this regard, policy makers should explore the possibility of conducting independent surveys of physicians and group practices in order to produce more timely data. Researchers surveying physicians and physician group practices could field their own data collection efforts and, at the same time, work with the National Center for Health Statistics to supplement the NAMCS sample and create additional survey modules. New surveys of physician group practices should start with a national random sample of physicians, or build off of an existing physician survey, and use this sample to create a sample of groups. Researchers could design a survey module for practice managers that include questions on practice size, region, multi or single specialty, multi- or single-site location and market integration. Researchers designing new hospital survey efforts should consider partnering with the AHA.

Physicians need to know their patients because there is something inherently personal about disease and illness. IT must be used in the service of a goal that is deeply human. Medical school curricula will have to be changed to prepare future physicians to use IT. Standardization is also important for the future of health IT. To achieve that goal, some authoritative source, consortia of leading businesses or probably the government, will have to set the standards.

Conclusions

Healthcare visionaries foresee a time when all types of health-related information exists electronically and can be reliably and securely accessed by any number of authorized parties and entities to improve the health of an individual, a specific community, or the U.S. population as a whole. The innovations of information and communication technology are crucial for facilitating reliable, comprehensive and quality clinical and healthcare services. The result of having HIT is that more patients are now better informed and they thus feel more equipped to participate in the intervention process.

References

[1] Valacich, J. & Schneider, C. (2009). *Information Systems Today: Managing the Digital World*. Upper Saddle River, NJ: Prentice Hall Press.

[2] Bushko, R. G. (2002). *Future of health technology*. Amsterdam, The Netherlands: IOS Press.

[3] Saarinen, K. & Aho, M. (2005). Does the implementation of a clinical information system decrease the time intensive care nurses spend on documentation of care? *Acta Anaesthesiologica Scandinavica*, 49(1), 62-65.

[4] Ruben, A., Aaron, C., Darrell, G. & Neil, P. (2008). Hospital characteristics associated with highly automated and usable clinical information systems in Texas, United States. *BMC Medical Informatics and Decision Making*, 8, 39.

[5] Garg, A. X., Adhikari, N. K. J., McDonald, H., Rosas -Arellano, M. P., Devereaux, P. J., Beyene, J. et al. (2005). Effects of computerized clinical decision support systems on practitioner performance and patient outcomes. *JAMA: the journal of the American Medical Association*, 293(10),1223 -1238.

[6] Ash, J. S., Gorman, P. N., Seshadri, V., and Hersh, W. R. (2004). Computerized physician order entry in US hospitals: results of a 2002 survey. *Journal of the American Medical Informatics Association*, 11(2), 95 -99.

[7] Pallin, D. J., Sullivan, A. F., Kaushal, R. & Camargo, C. A. (2010). Health information technology in US emergency departments. *International Journal of Emergency Medicine*, 3(3), 181-185.

[8] Hanuscak, T. L., Szeinbach, S. L., Seoane-Vazquez, E., Reichert, B. J. & McCluskey, C. F. (2009). Evaluation of causes and frequency of medication errors during information technology downtime. *American Journal of Health-System Pharmacy*, 66 (12), 1119-1124.

[9] Darbyshire, P. (200). Rage against the machine?: nurses' and midwives' experiences of using Computerized Patient Information Systems for clinical information. *Journal of clinical nursing*, 13(1), 17-25.

[10] Evans, J. M. M., Guthrie, B., Pagliari, C., Greene, A., Morris, A.D., Cunningham, S. et al. (2008). Do general practice characteristics influence uptake of an information technology (IT) innovation in primary care? *Informatics in Primary Care*, 16(1), 3-8.

[11] Sorby, I. D., Nytro, O., Tveit, A. & Vedvik, E. (2005). Physicians' Use of Clinical Information Systems in the Discharge Process: An Observational Study. *Studies in health technology and informatics*, 116, 843-848.

[12] Wulsin, L., Dougherty, A. (2008). *Health Information Technology-Electronic Health Records: A Primer*. California State Library. A report by the California Research Bureau.

[13] Davidson, C. M. & Santorelli, M. J. (2009). *Barriers to Broadband Adoption: A Report to the Federal Communications Commission.* A report of the Advanced Communications Law & Policy Institute, New York Law School.

[14] Hillestad, R., Bigelow, J., Bower, A., Girosi, F., Meili, R., Scoville, R. et al. (2005). Can electronic medical record systems transform health care? Potential health benefits, savings, and costs. *Health Affairs*, 24(5), 1103-1117.

[15] Zhou, W. & Piramuthu, S. (2010). Framework, strategy and evaluation of health care processes with RFID. *Decision Support Systems*, 50, 222-233.

[16] Meiller, Y., Bureau, S., Zhou, W. & Piramuthu, S. (2011). Adaptive knowledge-based system for health care applications with RFID-generated information. *Decision Support Systems*, 51, 198-207.

[17] Kuperman, G. J. & Gibson, R. F. (2003). Computer physician order entry: benefits, costs, and issues. *Annals of internal medicine*, 139(1), 31-39.

[18] Litan, R. E. (2008). *Vital signs via broadband: remote health monitoring transmits savings, enhances lives.* A report by Better Health Care Together.

[19] Meyer, M., Kobb, R. & Ryan, P. (2002). Virtually healthy: Chronic disease management in the home. *Disease Management*, 5(2), 87-94.

[20] Litan, R. E. (2005). *Great expectations: Potential economic benefits to the nation from accelerated broadband deployment to older Americans and Americans with disabilities.* A report by the New Millennium Research Council.

[21] Dilmaghani, R. B., Manoj, B. S. & Rao, R. R. (2006). Emergency communication challenges and privacy. *Proceedings of the 3rd International ISCRAM conference*. Newark, NJ.

[22] Hill, E. (2007). *A State Policy Approach: Promoting Health Information Technology in California*. An LAO report.

[23] Hall, C. (2009). *What is a Health Information Exchange?* A report by The College of American Pathologists.

[24] Orszag, P. R. (2008). *Evidence on the Costs and Benefits of Health Information Technology*. A report by the Congressional Budget Office.

[25] Girosi, F., Meili, R. & Scoville, R. (2009). *Extrapolating Evidence of Health Information Technology Savings and Costs*. Retrieved from: http://www.rand.org/pubs/monographs/MG410.html.

[26] McGee, M. K. (2011). Healthcare IT Priorities: A Singlular Focus. *Information Week: Analytics*. Retrieved from http://reports.informationweek.com/abstract/105/5934/healthcare/research-healthcare-it-priorities-survey.html

[27] Condon, S. (2009, March 18) *Health care experts warn of challenges for IT adoption*. CNET News.

[28] Dolan, P. (2010, November 22) *EMR adoption rates up, with small practices left behind*. amednews.com.

[29] Bansal, G., Zahedi, F. & Gefen, D. (2010). The impact of personal dispositions on information sensitivity, privacy concern and trust in disclosing health information online. *Decision Support Systems*. 49, 138-150.

A LINEARIZED GERBER FATIGUE MODEL

Edward E. Osakue, Texas Southern University

Abstract

A new bending fatigue design model based on the linearization of the Gerber failure criterion is presented here. The model divides the design diagram into two failure regimes: dynamic fatigue and static fatigue. Fatigue failure is generally of a brittle fracture nature. Dynamic fatigue failure is associated with fatigue strength of a material, while static fatigue failure is associated with the ultimate tensile strength of a material. Some examples are presented in order to demonstrate the use of the model in design verification and sizing. This model provides solutions that are slightly more conservative than the Gerber criterion but not as conservative as the Goodman criterion. This means that when bending fatigue is a significant serviceability criterion, it will yield design solutions with smaller sizes, giving more cost-effective designs. Generally, smaller components weigh less and are often cheaper to manufacture, enhancing profitability. Also, smaller products will reduce the rate of depletion of scarce materials. This new model, then, has the potential to help achieve economical designs for machine and structural members.

Introduction

Many machine and structural members are loaded by repeated or cyclic forces that can lead to fatigue failure. About 80% to 90% of the failures of machine and structural members result from fatigue [1], [2]. Hence, fatigue failure represents a significant proportion of failure problems in mechanical and structural systems. The importance of fatigue failure was first recognized by Pencelot in 1839. Rankine made similar observations about fatigue phenomena in 1843. Between 1843 and 1870, Wholer designed fatigue testing machines and used them to conduct many fatigue tests, including the investigation of the influence of stress concentration due to changes in cross sections [4]. Fatigue failure since then has been studied by numerous scientists and engineers such as Gerber, Goodman, Soderberg, Miner, Petersen, Marin, and a host of others [3-5]. Norton, [3] provides a time line and summary of many contributors in fatigue science and technology.

Several approaches are available for fatigue design and analysis [3]. The focus in this study is the stress-life (*S-N*) approach. An *S-N* diagram [6] displays three distinct portions judging by its slope. Hence, in the *S-N* approach, the fatigue load cycles may be divided into low-cycle fatigue, high-cycle fatigue and infinite-life fatigue regimes. However, there is no universal agreement on the dividing line between these regimes as overlap exists from classification by different authors [3], [7], [8]. Low-cycle fatigue is generally in the range of 1 to 10^3 load cycles and high-cycle fatigue is between 10^3 and 10^7 load cycles. Infinite-life fatigue is generally 10^6 load cycles and above. In low- and high-cycle fatigue, the life of a component is measured as the number of load cycles before failure. In infinite-life fatigue, the material is able to sustain an unlimited number of load cycles at some low stress levels. For most steel materials, infinite-life is observable between $2x10^6$ and 10^7 load cycles. For materials without apparent infinite life, it is often taken to be 10^8 or $5x10^8$ cycles.

The stress state in bending fatigue is appraised from the maximum and minimum stress values imposed on the structural or machine member during one load cycle. The exact variation of the stress during the cycle does not seem to be particularly relevant [1], [7]. The damage from a fluctuating bending stress state is assessed on the basis of the mean and alternating stress components. The alternating and mean stress components (please refer to Appendix for nomenclature) per cycle are, respectively:

$$\sigma_a = \frac{1}{2}(\sigma_{\max} - \sigma_{\min}) \tag{1}$$

$$\sigma_m = \frac{1}{2}(\sigma_{\max} + \sigma_{\min}) \tag{2}$$

When σ_m is positive or tensile during a fatigue load cycle, the material can fail at stress levels lower than the yield strength. Several models [3], [4], [9-12] are available in that address the influence of tensile mean stress on fatigue life. Among these are the Gerber (Germany, 1874), Goodman (England, 1899), and Soderberg (USA, 1930) models. According to Norton [3], the Gerber criterion is a measure of the average behavior of ductile materials in fatigue resistance, while the modified Goodman criterion is that of minimum behavior. Shingley and Mischke [10], state that the Gerber criterion falls centrally on experimental data while the modified Goodman criterion does not. The modified Goodman criterion is often used as a design criterion because it is more conservative than the Gerber criterion. Also, the modified Goodman criterion is simpler in application, especially in determining the size of members due to its linear nature. The use of the Gerber criterion in the deter-

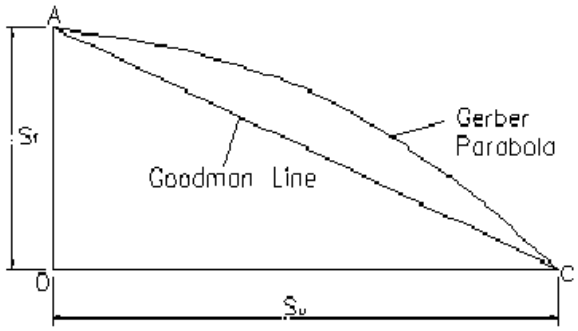
mination of member size is generally more computationally intensive and so rather unattractive for many designers. If the Gerber criterion is linearized, it can be used to determine the size of machine and structural members like the modified Goodman criterion.

The objective of this study was to develop a linearized model of the Gerber criterion so that it could be used in design sizing, like the modified Goldman criterion, without iterations. Because the Gerber criterion represents average behavior of ductile materials, one can expect 50% reliability. Hence, using this rule in a probabilistic model for design sizing means definite probability goals can be achieved. Over-design can be avoided by using probabilistic methods, while still ensuring the safety of a component [13]. Because it is less conservative than the modified Goodman criterion, a linearized Gerber rule will lead to reduced component sizes so that designs can be more cost effective as smaller components are lighter and often easier to make. In a global and technologically advancing world economy, cost-effective designs are a competitive edge. Lastly, material usage per product will be reduced, which will help to conserve scarce resources.

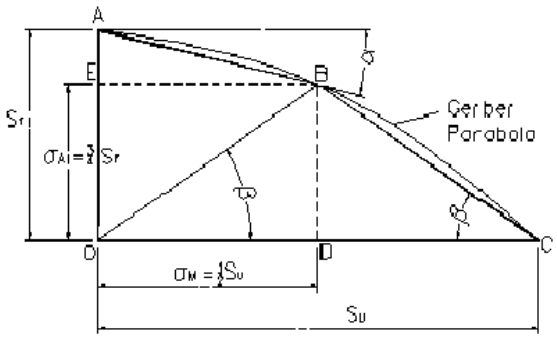
Linearizing the Gerber Criterion

Figure 1a shows a bending fatigue diagram with tensile mean stress indicating the Gerber parabola and the modified Goodman line. In this figure, the fatigue strength of the material is on the vertical axis and the ultimate tensile strength is on the horizontal axis. The Soderberg model is not shown because it is said to be more conservative than the modified Goodman rule [10] and is seldom used.

The basic idea of a linearized Gerber criterion is the approximation of the Gerber parabola with two line segments. Figure 1b shows two line segments AB and BC as approximations of the Gerber curve. This effectively divides the allowable design space into two triangles OAB and OBC with line OB common between these two triangles. In region OAB, material failure will most likely result from the predominant influence of the alternating stress and is called the dynamic fatigue regime. Brittle fracture is expected to be the dominant mode of failure in this regime. The failure line in triangle OAB is line AB and it makes angle α with the horizontal line. In region OBC, material failure will most likely result from the predominant influence of the mean stress and is called the static fatigue regime. Brittle fatigue failure is still expected in this regime, but some type of yielding is conceivable, especially at the micro-level, before gross fracture. The failure line in triangle OBC is line BC and it makes angle β with the horizontal line. It can be seen that lines AB and BC are on the conservative side of the Gerber parabola. The linearized Gerber model is defined by angles α (or ψ) and β (or η_t).



a) Common



b) Linearized Gerber

Figure 1. Bending Fatigue Design Diagrams

The Gerber parabola in bending fatigue is described by the equation:

$$\sigma_A = S_f \left[1 - \left(\frac{\sigma_M}{S_u} \right)^2 \right] \tag{3}$$

Referring to Figure 1b, when $\sigma_M = \frac{S_u}{2}$ in Equation (3), then:

$$\sigma_A = \frac{3}{4} S_f \tag{4}$$

Therefore,

$$\tan\alpha = \psi = \frac{OA - OE}{OD} = \frac{S_f - \sigma_A}{\frac{S_u}{2}} = \frac{S_f}{2S_u} = \frac{\psi_s}{2} \tag{5}$$

$$\tan\beta = \eta_t = \frac{OE}{OD} = \frac{3S_f}{2S_u} = \frac{3\psi_s}{2} \tag{6}$$

Equations (5) and (6) depend on ψ_s , which is obtained from the basic fatigue ratio ψ_o. ψ_o is based on standard polished laboratory specimens. ψ_s is obtained by multiplying ψ_o by adjustment factors such as temperature, size or reliability [3], [9], [10], depending on service conditions. Some representative values of ψ_o are wrought steel at 0.5; cast steel, nodular cast iron, aluminum and copper alloys at 0.4; gray cast iron at 0.35; and normalized nodular cast iron at 0.33 [3], [10]. Some scatter can be expected with fatigue ratio data. Also, at relatively high ultimate tensile strengths, the fatigue ratio drops, so care is needed in using these values. Note that ψ_s will normally be smaller than ψ_o in service conditions.

Effective Bending Stress

With the angles α and β now known, the effective bending stress, resulting from a combination of alternating stress and mean stress, needs to be determined. Any combination of these stresses will have a load line that passes through the origin with a slope given by:

$$\eta = \frac{k_\sigma \sigma_a}{\sigma_m} \tag{7}$$

Equation (7) has the stress concentration factor k_σ applied to the alternating normal stress. This is necessary for realistic estimates of stresses at cross-sections with notches. According to Collins et al. [9], experimental studies indicate that stress concentration factors should be applied only to alternating components of stress for ductile materials in fatigue loading. However, stress concentration factors should be applied to both alternating and mean stresses in brittle materials when loaded in fatigue. Now, the load line factor η determines the fatigue failure regime that is appropriate for a particular situation. If η is equal to or greater than the load line transition factor, η_t , then the design point will be inside triangle OAB in Figure 1b, and the dynamic fatigue failure regime would apply. If η is less than η_t , the design point will be inside triangle OBC in Figure 1b, and the static fatigue failure regime would apply.

Dynamic Fatigue Failure Regime: $\eta \geq \eta_t$

Figure 2a shows a fatigue bending stress state with a load line in the dynamic fatigue failure regime. The effective bending stress is represented by OF and alternating stress by OE. OD represents the mean, or steady, stress. Note that lines FG and AB, the failure lines, are parallel. This ensures that effective stress is being mapped with the appropriate failure rule. If these two lines are not parallel, a different failure criterion would apply to line FG, introducing distortion to the failure rule.

Referring to Figure 2a:

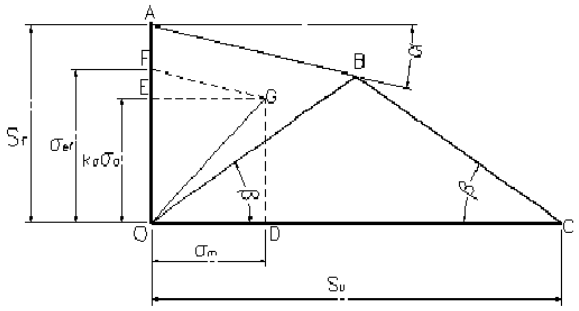
$$OF = \sigma_{ef} \qquad OE = k_a \sigma_a \qquad OD = EG = \sigma_m$$

$$EF = EG \tan\alpha = \psi \sigma_m \tag{8}$$

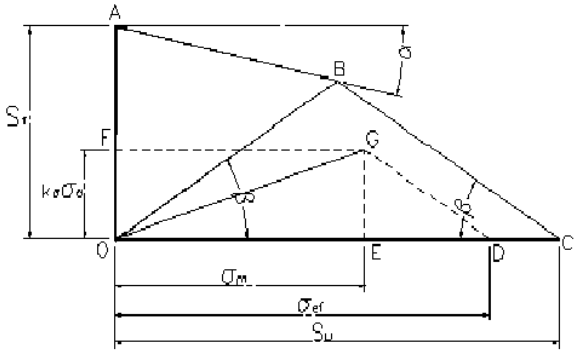
$$OF = OE + EF \tag{9}$$

$$\sigma_{ef} = k_a \sigma_a + \psi \sigma_m = k_a \sigma_a + \frac{1}{2} \psi_s \sigma_m \tag{10}$$

Equation (10) gives the effective normal stress for the new model in the dynamic fatigue failure regime.



a) **Dynamic Fatigue Failure Regime**



b) Static Fatigue Failure Regime

Figure 2. Effective Bending Stresses in Failure Regimes

The effective bending stress in this situation is projected on the fatigue strength (vertical) axis. Fatigue strength is a dynamic material property. The effective stress based on the modified Goodman model can be expressed as:

$$\sigma_{ef} = k_a \sigma_a + \frac{S_f}{S_u} \psi \sigma_m = k_a \sigma_a + \psi_s \sigma_m \tag{11}$$

The effective normal stress based on the Gerber model [9] is:

$$\sigma_{ef} = \frac{k_\sigma \sigma_a}{1 - \left(\frac{\sigma_m}{S_u}\right)^2} \tag{12}$$

The factor of safety in the dynamic fatigue failure regime is obtained as:

$$n_s = \frac{S_f}{\sigma_{ef}} \tag{13}$$

Static Fatigue Failure Regime: $\eta < \eta_t$

Figure 2b shows a fatigue bending stress state with a load line in the static fatigue failure regime. In this regime, the effective normal stress is projected on the horizontal axis. The effective bending stress is represented by OD and alternating stress by OF. Note that lines DG and BC are parallel, which makes the slope of line DG equal to that of line BC, the failure line. As in the static fatigue failure regime, this ensures that effective stress is being mapped with the appropriate failure rule.

Referring to Figure 2b:

$$OD = \sigma_{ef}$$

$$EG = OF = k_a \sigma_a$$

$$FG = OE = \sigma_m$$

$$\tan\beta = \eta_t = \frac{OF}{ED} \qquad ED = \frac{OF}{\tan\beta} = \frac{k_\sigma \sigma_a}{\eta_t} = \frac{2k_\sigma \sigma_a}{3\psi_s} \tag{14a}$$

$$OD = OE + ED \tag{14b}$$

Thus:

$$\sigma_{ef} = \sigma_m + \frac{k_\sigma \sigma_a}{\eta_t} = \sigma_m + \frac{2k_\sigma \sigma_a}{3\psi_s} \tag{15}$$

Equation (15) gives the effective mean normal stress by the new model for the static fatigue failure regime. The effective bending stress in this situation is projected on the tensile strength (horizontal) axis. Tensile strength is a "static" material property. The effective stress based on the modified Goodman model can be expressed as:

$$\sigma_{ef} = \sigma_m + \frac{k_\sigma \sigma_a}{\psi_s} \tag{16}$$

An expression for the Gerber effective mean normal stress can be found by replacing S_u with σ_{ef} in the Gerber criterion and simplifying it. That is,

$$\sigma_{ef} = \frac{\sigma_m}{\sqrt{1 - \frac{k_\sigma \sigma_a}{S_f}}} \tag{17}$$

The factor of safety in the static fatigue failure regime is obtained as:

$$n_s = \frac{S_u}{\sigma_{ef}} \tag{18}$$

Avoiding Yield at the First Load Cycle

There is a possibility that when a static fatigue failure condition exists, a member may yield at the first load circle [3], [10]. The Langer or the yield line shown in Figure 3 is inclined at 45° to the horizontal. If a stress state lies to the right of line DF, then yielding would have occurred. If this happens, local yielding can occur, which can lead to changes in straightness and strength (local strain hardening is also possible), resulting in unpredictable loading [10]. The line DE in Figure 3 is parallel to the failure line BC. Now, if the angle β is smaller than 45°, then yielding can be prevented in the static fatigue failure regime by translating line BC to the position of line DE, which does not change the failure criterion. For most materials used in fatigue design, β will be smaller than 45° since the high value of ψ_o is about 0.5 as was indicated previously. This limits the angle β from Equation (7) to about 37° in a worst-case scenario. However, ψ_s is normally smaller than ψ_o , so the value of β in service will even be lower than 37°. To translate line BC to position DE means the safety factor, n_s , on the ultimate strength should be sufficiently high to preclude yielding. This condition is satisfied when

$$n_s \geq n_o = \frac{S_u}{S_Y} \tag{19}$$

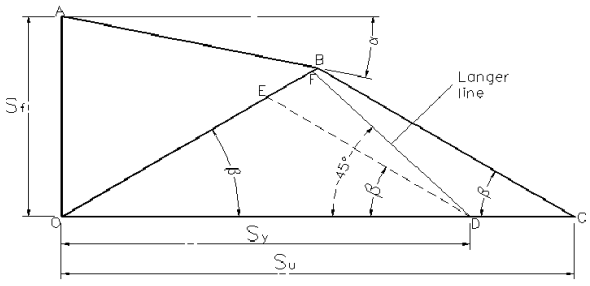


Figure 3. Avoiding Yield at First Load Cycle

Design Sizing Application

Bending stress can be expressed as a function of the bending moment and section modulus:

$$\sigma_a = \frac{M_a}{Z} \tag{20}$$

$$\sigma_m = \frac{M_m}{Z} \tag{21}$$

$$\eta = \frac{k_\sigma \sigma_a}{\sigma_m} = \frac{k_\sigma M_a}{M_m} \tag{22}$$

Dynamic Fatigue Failure Design: $\eta \geq \eta_t$

For design sizing applications in a dynamic fatigue failure regime, the task is to find a Z value that will satisfy the strength serviceability criterion. This condition is expressed as:

$$\sigma_{ef} \leq \frac{S_f}{n_s} \tag{23}$$

From Equations 10, 20, 21 and 23:

$$\sigma_{ef} = k_a \sigma_a + \psi \sigma_m = \frac{k_\sigma M_a}{Z} + \frac{\psi_s M_m}{2Z} \leq \frac{S_f}{n_s} \tag{24}$$

So that:

$$Z \geq \frac{n_s}{S_f}\left\{k_\sigma M_a + \frac{1}{2}\psi_s M_m\right\} \tag{25}$$

Static Fatigue Failure Design: $\eta < \eta_t$

For design sizing applications in a static fatigue failure regime, the task is to find a Z value that will satisfy the strength serviceability criterion:

$$\sigma_{ef} \leq \frac{S_u}{n_s} \tag{26}$$

Combining Equations 13, 20, 21 and 26:

$$Z \geq \frac{n_s}{S_u}\left\{\frac{k_\sigma M_a}{\eta_t} + M_m\right\} = \frac{n_s}{S_u}\left\{\frac{2k_\sigma M_a}{3\psi_s} + M_m\right\} \tag{27}$$

The section modulus, Z , depends on the shape and dimensions of the shape of the cross-section of a member. For simple shapes such as circles and rectangles, the formula for Z is available in structural and machine design textbooks. In structural design, values of Z can be obtained from tables for structural steel shapes; for example, AISC Steel Construction Manual.

Some Applications of the Linearized Gerber Model

The linearized Gerber model (LGM) was applied in three examples. The first is a case of possible dynamic fatigue failure taken from Norton [3], while the second example is a case of possible static fatigue failure and a modification of the first example. This example was used because it is described as a typical design problem [3]. The model application in these examples is that of design verification in which the adequacy of a design is assessed on the basis of a factor of safety for a member with a known form or 3D figure. A design is accepted as adequate if the factor of safety is at least equal to a desired value. A factor of safety greater than unity is necessary for failure avoidance. Design verification is a task often performed in the detail/prototype phase of a design project. The third example is a redesign of the components of the first example, demonstrating the application of the new model in design sizing. The task in design sizing is choosing the form and determining the size of a member for a desired factor of safety.

The form of a member is defined by its length, cross-sectional shape and dimensions over its length. In general, the cross-section may vary along the length of a member, but this makes analysis more complicated and costly. Constant cross-sectional members are usually the first choice, especially during preliminary design, but modifications often occur later in the design process. The length of a member is often based on space limitation and may be estimated in a preliminary layout diagram but can be refined later, taking into consideration rigidity and strength. The cross-section can be sized for an assumed shape based on fatigue strength or other serviceability criteria. Design sizing is a task often performed in the preliminary phase of a design project.

Example 1

Figure 4 shows one of two brackets attached to a machine frame. The brackets carry a combined fluctuating load varying from a minimum of 890N to a maximum of 9,786N, (data converted to SI Units) [3]. The load is shared equally by the brackets; the maximum allowable lateral deflection was 0.51mm for each bracket, each of which should be designed for 10^9 load cycles. The load-time function was sinusoidal, maximum cantilever length was 152mm, and the

operating temperature was 50°C. Trial dimensions were b = 51mm, h = 25.4mm, H = 28.6mm, r = 12.7mm and l = 127mm. The brackets were machined to size from stocks. Norton [3] recommends values of k_σ = 1.16 and Z = 5463.45mm^3. The brackets were made from SAE 1040 steel with S_u = 550MPa, S_y = 414MPa and S_f = 150MPa at 99.9% reliability.

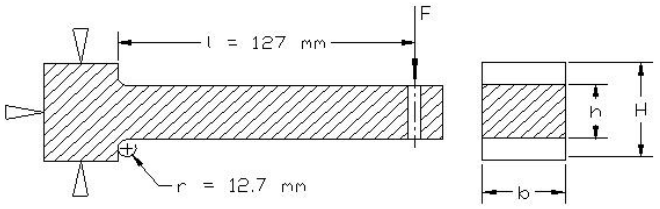


Figure 4. Sample Bracket

Solution 1

The expected 10^9 load cycles for the brackets was in the range of infinite-life regime. This simplified the estimation of the service fatigue strength [3]. Now S_f = 150MPa because it was evaluated at 99.9% reliability. At 50% reliability, the reliability adjustment factor was 1.0; at 99.9% reliability, it was 0.753 [3]. Because the Gerber model is said to represent average behavior of materials, it was consistent to use the model at a 50% reliability level. Hence, the fatigue strength at 50% reliability was S_f = 200MPa.

Equations 7, 10-13, 20 and 21 were coded in Excel, Microsoft spreadsheet software for analysis in dynamic fatigue bending fatigue failure. Table 1 shows the layout of the Excel page. It consists of two main sections of Input and Output. The maximum bending moments and other data were provided as input data and the codes generated the output data. The same data were used to run the codes developed for the static fatigue failure mode. The critical section of the bracket is at the fillet position where the bending moment is maximum. The maximum bending moments were evaluated to be 339,280.5Nmm for M_m and 282,130.5Nmm for M_a.

Table 2 summarizes the safety factor results for the design verification for Example 1. For the same design conditions, a more conservative model will give a smaller safety factor. From Table 2, the LGM was observed to be slightly more conservative than the Gerber model but less conservative than the modified Goodman model. This shows that the new model is an improvement on the modified Goodman criterion and, thus, will help to conserve material resources if used. The Gerber model gives a safety factor of 7.18 in static fatigue failure analysis, indicating that dynamic fatigue failure is actually more likely (lower safety factor value). The modified Goodman model shows no such discrimination. Certainly, the new model classification into dynamic fatigue and static fatigue failure regimes seems realistic.

Table 1. Calculations for Example 1

DESIGN VERIFICATION: DYNAMIC FATIGUE		
INPUT DATA		
Material Properties		
Tensile Strength (MPa), S_u	550	
Yield Strength (MPa), S_y	414	
Fatigue Strength (MPa), S_f	200.0	
Design Factors		
Overload factor	1.00	
Stress concentration factor, k_σ	1.160	
Bending Moments		
Alternating moment (Nmm), M_a	282448.0	
Mean moment (Nmm), M_m	338938.0	
Section modulus (mm^3), Z	5463.45	
OUTPUT DATA		
Service fatigue ratio: ψ_s		0.3636
Alternating stress (MPa), σ_a		59.9694
Mean stress (MPa), σ_m		62.0374
Load line transition factor, η_t		0.5455
Load line factor, η		0.9667
Linearized Gerber Model		
Model effective stress (MPa), σ_{ef}		71.249
Design safety factor, n_s		***2.81***
Modified Goodman Model		
Model effective stress (MPa), σ_{ef}		82.528
Design safety factor, n_s		***2.42***
Gerber Model		
Model effective stress (MPa), σ_{ef}		60.742
Design safety factor, n_s		***3.30***

Example 2

The problem of Example 1 was analyzed with a fluctuating load on a bracket varying from a minimum of 3,114N to a maximum of 4,893N. Other factors in the problem remained unchanged.

Table 2. Models of Comparison for the Dynamic Fatigue Failure Design

Model	Dynamic Fatigue Failure		Static Fatigue Failure	
	Safety Factor	% Difference	Safety Factor	% Difference
Gerber	3.30	0	7.18	0
LGM	2.81	14.85	Not applicable	
Modified Goodman	2.42	26.67	2.43	66.16

Solution 2

As in the previous problem, Equations 7 and 15-21 were coded in Excel for analysis in static fatigue bending fatigue failure. This Excel page was similar to that of Table 1 and, thus, not shown here. The maximum bending moments and other input data were supplied to the codes for both failure modes. The critical section is at the fillet location as in Example 1. The maximum bending moments were determined to be 508,508Nmm for M_m and = 112,966.5Nmm for M_a.

Table 3. Models of Comparison for the Static Fatigue Failure Design

Model	Dynamic Fatigue Failure		Static Fatigue Failure	
	Safety Factor	% Difference	Safety Factor	% Difference
Gerber	8.1	0	5.50	0
LGM	Not applicable		4.02	27.6
Modified Goodman	3.46	57.28	3.47	37.8

Table 3 summarizes the results for design verification for Example 2. The minimum safety factor, n_o , to avoid yield at the first load cycle was 1.33. From Table 3, the LGM was again observed to be slightly more conservative than the Gerber model but less conservative than the modified Goodman model. The safety factor for the Gerber model in the static fatigue failure regime was smaller than that of the dynamic fatigue failure regime, indicating that static fatigue failure is more likely. The 27.6% deviation of LGM from the Gerber model can be explained by taking a closer look at Figure 1b. The maximum deviation of line BC from the Gerber parabola is expected about its midpoint. Since the transition load line angle is $\beta = 28.54^o$ and the load line angle is 14. 56^o (about half of β) in this example, 27.6% deviation represents about the maximum error to be expected from the new model for this design condition. Again, the modified Goodman model showed no such difference in safety factor due to its single linear relationship. Therefore, the new model classification into dynamic fatigue and static fatigue failure regimes appear to be very realistic.

Example 3

Example 3 was a redesign of the brackets of Figure 4 such that b was half h , and where h and H maintain the same ratio for a minimum safety factor of 2.5. The material and other conditions remained the same as stated in Example 1.

Solution 3

Equations 20-22, 25 and 27 were coded in Excel for design sizing for the new model. The task in this problem was to determine the section modulus Z for the critical section which was at the fillet location in Figure 4. From the section modulus value, the dimensions of the cross-sectional shape can be determined once the shape type was chosen. The shape of the cross-section was rectangular, as shown on the right side of Figure 4. The shape factor was taken to be the ratio of section width to section height, which was 0.5 in this example. From Example 1, M_m = 339,280.5Nmm and M_a = 282,130.5Nmm. Now, n_s = 2.5 and k_σ was taken as 1.3 as a trial value.

Table 4 shows the layout of a portion of the Excel pages that made up the codes in Excel. The full layout consisted of three sections of Input, Processing and Output. The processing page is not shown in Table 4 and only a portion of the output section is shown. The parameters used in the developed equations are indicated. The bending moments evaluated previously along with other data were provided as input and the codes generated the output data. The section depth, h, was calculated to be 40.06mm but was chosen to be 40mm; the width was chosen to be 20mm. With these dimensions:

$$H = 1.125x40 = 45\ mm; \qquad \frac{H}{h} = \frac{45}{40} = 1.125$$

$$\frac{r}{h} = \frac{12.7}{40} = 0.3175$$

Based on the ratios of 1.125 and 0.3175, and from Figure 4.36 in the book by Norton [3], k'_σ was read to be 1.3 and k_σ was evaluated to be 1.2627. k'_σ and k_σ are related by the notch sensitivity factor. Using the same procedure as in Example 1, the design safety factor, n_s was evaluated to be 2.55 for the new model and 2.95 for the Gerber model. These values are relatively close to the desired value of 2.5. The deflection at the point of load application was calculated to be 0.151mm and 0.196mm at the end of the bracket. These values were much lower than the maximum allow-

able value of 0.51mm. The cross-sectional area of the old bar was 1290mm^2. The new bar had a cross-sectional area of 800mm^2. This is a 38% reduction in area and consequently a 38% reduction in weight or material cost at 36.4% of maximum allowable deflection. A 38% reduction in material cost could translate into thousands if not millions of dollars in savings in a large volume production!

Table 4. Calculations for Example 3

DESIGN SIZING: DYNAMIC FATIGUE FAILURE	
INPUT DATA	
Material Properties	
Tensile Strength (MPa), S_u	550
Yield Strength (MPa), S_y	414
Fatigue Strength (MPa), S_f	200.00
Factors	
Safety factor, n_s	2.50
Overload factor	1.00
Stress concentration Factor, k_σ	1.3000
Bending Moments	
Alternating moment (Nmm), M_a	282448
Mean moment (Nmm), M_m	338938
Section Shape	
Rectangular: shape factor	0.50
OUTPUT DATA	
Failure Type Assessment	
Load line transition factor, η_t	1.0500
Load line factor, η	0.5691
Failure Type (DFF* or SFF*)	DFF
Dynamic Fatigue Design	
Rectangle	
Chosen Dimensions	
Height (mm), h	40.00
Width (mm), b	20.00
Section modulus (mm^3), $Z = \frac{bh^2}{6}$	5333.33

*****DFF = Dynamic Fatigue Failure
*SFF = Static Fatigue Failure.

Conclusion

A linearized Gerber model (LGM) was developed in order to simplify the use of the Gerber model in design sizing. The model divides the fatigue design diagram into two portions of dynamic fatigue failure and static fatigue failure regimes. As shown in Examples 1 and 2, this division correctly identifies the more likely mode of failure in design situations. The linearized model is less computationally intensive than the Gerber model and less conservative than the modified Goodman model. It defines a minimum safety factor for static fatigue failure design if yielding must be precluded.

References

[1] Kravchenko, P. Y. (1964). *Fatigue Resistance*. Oxford Pergamon, New York.

[2] Sachs, N. (1999, August). Root Cause Failure Analysis Interpretation of Fatigue Failures. *Reliability Magazine*.

[3] Norton, R. L. (2000). *Machine Design: An Integrated Approach*. Prentice-Hall, Upper Saddle River, New Jersey.

[4] Hidgon, A., Ohseen, E. H., Stiles, W. B. & Weesa, J. A. (1967). *Mechanics of Materials*. (2nd ed.). Wiley, New York.

[5] Polak, P. (1976). *A background to Engineering Design*. MacMillan, London.

[6] EPI Inc., (2008), "Metal Fatigue-Why Metal Parts Fail from Repeatedly Applied Loads".http://www.epi-eng.com/mechanical_engineering_basics/fatigue_in_metals.ht.

[7] Dieter, E. G. (1976). *Mechanical Metallurgy*. (2nd ed.). McGraw-Hill, New York.

[8] Shigley, J. E. & Mitchell, L. D. (1983). *Mechanical Engineering Design*. McGraw-Hill, New York.

[9] Collins, A. J., Busby, H. & Staab, G. (2010). *Mechanical Design of Machine Elements*. John Wiley & Sons, New Jersey.

[10] Shigley, J. E. & Mischke, C. R. (1996). *Standard Handbook of Machine Design*. (2nd ed.). McGraw-Hill, New York.

[11] Spotts, M. F. (1985). *Design of Machine Elements*. Prentice-Hall, Englewood Cliffs.

[12] Juvinall, R. C. (1983). *Fundamentals of Machine Components Design*. Wiley, New York.

[13] "Understanding Probabilistic Design", http://www.kxcad.net/ansys/ANSYS/ansyshelp/Hlp_G_ADVPDS1.htm

Appendix: Nomenclature

α = dynamic fatigue failure angle in fatigue

ψ = dynamic fatigue failure slope factor

$\psi_s = \frac{S_f}{S_u}$ = service fatigue ratio

$\psi_o = \frac{S_f^{'}}{S_u}$ = basic fatigue ratio

β = load line transition angle in fatigue

η = load line slope factor in fatigue

η_t = load line slope transition factor in fatigue

σ_A = Gerber alternating failure stress

σ_M = Gerber mean failure stress

σ_a = nominal normal alternating stress

σ_m = nominal normal mean stress

σ_{ef} = effective bending stress

$k_\sigma^{'}$ = theoretical bending stress concentration factor

k_a = service bending stress concentration factor

$k_a = q(k_\sigma^{'} - 1) + 1$

q = notch sensitivity factor

h = depth of rectangular cross-section

b = width of rectangular cross-section

M_a = alternating bending moment

M_m = mean bending moment

n_s = factor of safety

n_o = minimum factor of safety to avoid yield at first load cycle

$S_f^{'}$ = fatigue strength of polished laboratory specimen

S_f = service fatigue strength

S_Y = yield strength

S_u = ultimate tensile strength

Z = section modulus of member

Biography

EDWARD E. OSAKUE is an assistant professor in the Department of Industrial Technology at Texas Southern University in Houston. He is a graduate faculty and the coordinator of the Design Technology concentration. He instructs students in engineering design, engineering graphics, and drafting. His research interests include economical design of mechanical and structural systems, low-velocity impact with friction, and effective curriculum delivery methods. Dr. Osakue can be reached at osakueee@tsu.edu

CHARGE CONDUCTION MECHANISM AND MODELING IN HIGH-K DIELECTRIC-BASED MOS CAPACITORS

Madhumita Chowdhury, The University of Toledo; Branden Long, The University of Toledo; Rashmi Jha, The University of Toledo; Vijay Devabhaktuni, The University of Toledo

Abstract

In this study, four MOS capacitors were fabricated (Al_2O_3/Ti, Al_2O_3/W, HfO_2/Ti, and HfO_2/W) on Si substrates. Temperature-dependent measurements were performed to study the gate leakage current. To investigate the metal/high-k dielectric interface, charge transport mechanisms of the gate leakage current were examined. I-V curves were fitted to three main conduction methods, i.e., Frenkel-Poole (F-P) emission, Schottky emission, and tunneling mechanism, at both low and high electric fields with barrier heights extracted. Furthermore, to improve device-to-device uniformity, simplify the process of data acquisition, and increase yield, a neural network was employed to model the I-V relationships across different temperature ranges for all four samples.

Introduction

Gate leakage current is becoming the bottleneck in the design of high-speed and low-power consumption devices due to continuous scaling of the dielectric thickness of SiO_2 [1]. The continuous reduction in the thickness of the gate dielectric (SiO_2) results in very high gate leakage current due to direct tunneling of electrons through the SiO_2. This would finally lead to high power consumption even when the device is off, i.e., the gate voltage is less than the threshold voltage [2]. To ensure that Moore's Law remains valid in the next decade, leading integrated circuit (IC) manufacturers are now making a breakthrough by replacing the long-standing poly-silicon gate and SiO_2 [3] with a metal gate like Ti, W, Pt and high-k dielectrics such as HfO_2, Al_2O_3 and TiO_2 , respectively [4], [5]. The reason for replacement of poly-silicon gates with metal gates is to avoid high threshold voltages which arise due to Fermi-level pinning along with degradation of channel carrier mobility [6-8]. Fermi-level pinning is nothing but an inability to entirely move the Fermi level (E_F) across an Si band gap [9], [10]. Metal-induced gap states (MIGS) generated at poly-Si/SiO_2 interfaces tail off very quickly into the SiO_2 and cause very little pinning, for which E_F can be shifted completely through the Si gap by varying the work function of the electrode. However, in poly-Si/HfO_2 interface MIGS are more in number and die less swiftly when compared with SiO_2. Hence, a larger alteration in work function would be needed to oscillate E_F across the Si band gap [11]. A high-k-based dielectric helps in maintaining the same capacitance as that of a SiO_2 dielectric but with a thinner material layer [12].

However, in order to replace the longstanding combination of poly-Si/SiO_2 material with a metal/high-k dielectric, it is imperative to fully comprehend the interface mechanism. This can be done by studying the mechanism of charge transportation of gate leakage current. In spite of significant research progress in this area, the mechanism of charge transport responsible for the gate leakage current and its dependence on the gate dielectrics and metal electrodes is not well understood. To address this problem, the authors evaluated the charge conduction mechanism of gate leakage current in different high-k-dielectric-based MOS capacitors with metal electrodes on top. In addition, temperature-dependent measurements were made to compare the charge transport mechanism of atomic layer deposited (ALD) HfO_2-based MOS capacitors with that of Al_2O_3-based MOS capacitors with Ti and W electrodes. I-V curves were fitted for different conduction mechanisms at different temperature and voltage ranges. A neural network method of modeling was employed to ease the duplication of the same sample by avoiding calculating parameters like effective mass of an electron and material-specific parameters, which reduce the cost of fabrication and increase the speed of data acquisition.

Background

The reduction of the gate dielectric thickness is one of the core reasons for increases in gate leakage current of MOS devices [13]. To minimize this gate leakage current, many high-k gate dielectrics have been recommended as substitutes for SiO_2 in MOS structures with effective oxide thicknesses (EOT) lower than 1.5nm [14]. To better study the interface, temperature-dependent measurements are among the most important methods for determining the charge transport mechanism of the gate leakage current. A temperature-dependent study was performed on Ta_2O_5 and TiO_2 films to determine their conduction mechanisms and to verify whether the gate leakage current is supportable at high temperatures for either of these high-k dielectrics. From studies it was found that in Ta_2O_5, I–V curves showed stronger temperature dependence than in TiO_2 samples [15],

[16]. The chief reason was a lower electron barrier height in Ta_2O_5, which resulted in Schottky emission to dominate the charge conduction mechanism. On the contrary, the TiO_2 sample demonstrated tunneling as the dominant conduction mechanism in the high-field region, and F-P conduction in the low-field region. In the literature, it was reported that, for an HfO_2 dielectric, a high work function metal (Pt) would be responsible for the F-P emission conduction mechanism [17]. This is because the Schottky barrier height would be larger than the energy level of the traps. However, for an Al electrode, the Schottky emission dominated the conduction mechanism of gate leakage current. On the other hand, it has been proved that Al_2O_3-dielectric-based MOS capacitors with Al electrodes showed the F-P emission to be the dominating charge transport mechanism at the Al/Al_2O_3 interface [18].

In spite of these studies, a rigorous comparison was needed to identify which conduction mechanism dominates at which voltage level and temperature range for a specific metal/dielectric interface. In this study, a temperature-dependent comparison was made at different voltage ranges to identify the dominant current conduction mechanism. In addition, a neural network model based on a Quasi-Newton algorithm was employed for ease of sample reproduction by avoiding extracting parameters like barrier height and effective mass from equations which are labor-intensive and time -consuming processes. Neural network's ability to learn quickly for building convincing solutions to unformulated problems, manage computationally expensive models, deliver fast interpolative analysis, and attain very precise functional relationships between data sets are its major advantages [19]. It is a well-established method for modeling various processes in the semiconductor industry such as molecular beam epitaxy and plasma-enhanced chemical vapor deposition [20]. In a review of the literature, modeling of semiconductor process device characteristics was done in both the forward and inverse directions [21]. A multilayer perceptron neural network (MLPNN) was used for development of the model. In the forward direction, data obtained from the characteristics of earlier fabrication processing points were used as input to a MLPNN, and the last characteristic values were modelled. On the other hand, for inverse modeling, final DC device characteristic measurements of the total wafer were used as input to an MLPNN, and in-process characteristic data were modelled. This method eliminates the necessity to statistically describe parametric deviation across a wafer. In this study, modeling of gate leakage current was implemented for reducing the non-uniformity in the fabrication process and collecting additional data without fabricating the samples again.

Methodology

This section is divided into two parts: the experimental processes and the neural network modeling approach. The first section deals with the fabrication of the MOS capacitors while the latter discusses the modeling method used to calculate the output current.

Fabrication Process

HfO_2 and Al_2O_3 dielectric films were deposited on p-type Si wafers by an Atomic Layer Deposition ALD process at 300°C. The thicknesses for both of the dielectrics were measured to be ~60Å using an ellipsometer. W and Ti metals (gate electrodes) of 1,000Å were deposited by RF sputtering and were patterned by the lift-off technique. The back contact was formed by depositing 1,000Å of Al on the backside of the samples by RF magnetron sputtering followed by rapid thermal annealing (RTA) in an N_2 environment at 600ºC for 5 minutes in order to achieve a low-resistance ohmic contact. The samples were probed in the Lakeshore cryogenic probe station and I-V characteristics were obtained by a Keithley 4200 Semiconductor Characterization System.

Neural Network Modeling

The modeling of the collected data (gate leakage current) was done using a feed-forward neural network, which consisted of an input layer, a hidden layer and an output layer, as shown in Figure 1. Each layer was comprised of several elements called neurons, where the input layer was a relay function, the hidden layer was a sigmoid function and the output layer was a linear function of hidden neurons. Each neuron in a layer had an input from a previous layer and a constant (or bias), while its output was forwarded to the next layer. The inputs and outputs of the neuron were multiplied by a factor called weights. This feed-forward neural network developed a model from the training data supplied. The network is said to be feed forward because each component/element in a layer receives inputs only from the components/elements in the previous layer. In this study, the modeling was done by covering the entire range of temperature points between 300K – 400K using a Quasi-Newton algorithm. First, the data set available from the experiment was randomized and then segregated into two sections, namely a training data set and a validation data set. Voltages (-4V to 4V) and temperatures (300K, 350K, 400K) were taken as the two inputs for the neural network and current (corresponding to the voltage range) as the output. Out of all the data points, 80% were taken to train the neural network and 20% for validating the results. The neural net-

work approach was taken for modeling because the same or additional results can be obtained again without revisiting the entire process of fabrication and testing. This model will increase uniformity in fabrication, simplify the data acquisition process and, hence, increase future yield.

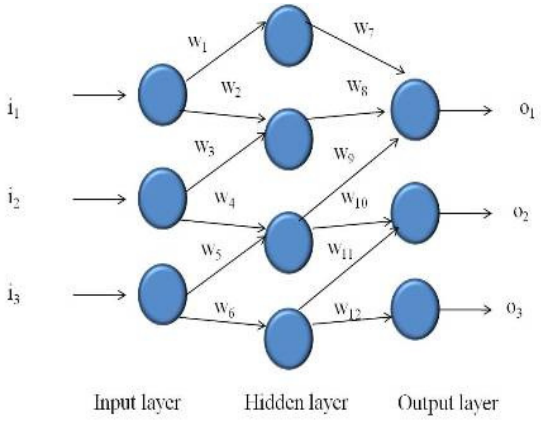


Figure 1. Neural Network Model

Results and Discussions

The room temperature measurements of gate leakage current for all four samples are shown in Figure 2. Currents for both devices with Ti electrodes resulted in much lower currents than those with W electrodes. To further understand their mechanism, temperature-dependent studies of gate leakage current were performed. Gate leakage currents were measured for temperatures varying from 120K to 400K (Figure 3), with an applied voltage of -4V. It was observed that at temperatures below 250K, current was almost constant. Hence, the tunneling mechanism dominates as the primary conduction mechanism for all devices in that temperature range (Figure 3). Above room temperature, any of the three methods—F-P, Schottky emission or tunneling mechanism—may dominate for the samples, depending upon the voltage range [22], [23]. Also, these processes may not be completely independent of each other. It can be seen from Figure 3 that at higher temperatures W-based samples show much-elevated current compared with those of Ti-based samples. For samples using W electrodes, it was observed that more than one mechanism was dominating at the same time. However, for samples with Ti electrodes, the tunneling mechanism was dominant at low temperatures.

I-V curves for all four samples were fitted at 300K, 350K, and 400K employing the following equations:

F-P Emission:

$$J \propto E \exp(-q(\phi_b - \sqrt{(qE/\pi\xi)}/KT) \qquad (1)$$

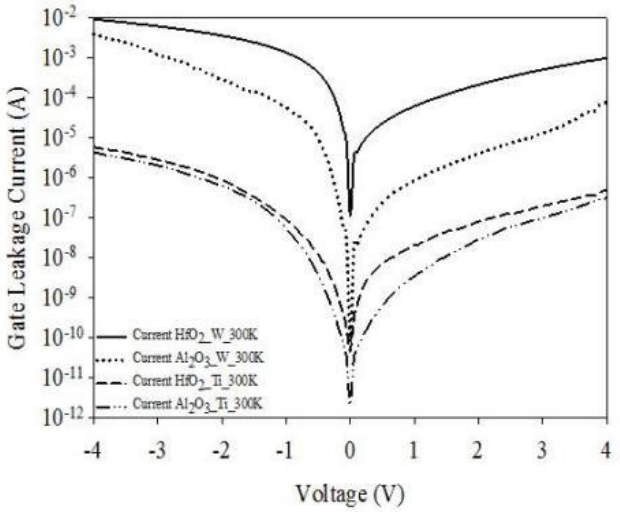


Figure 2. I-V Measurements for Four Samples at Room Temperature

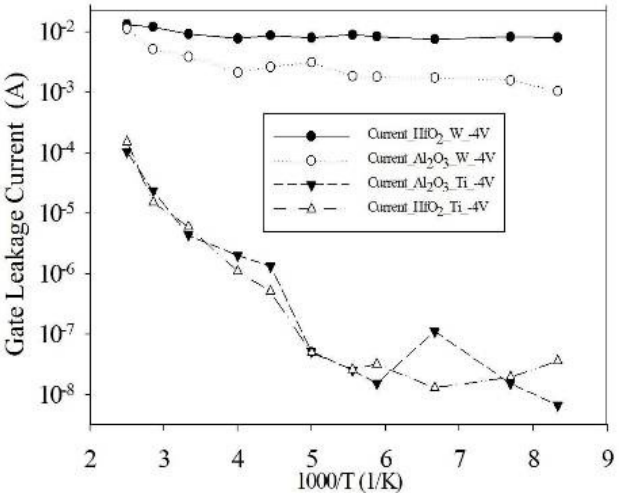


Figure 3. Gate Leakage Current vs Temperature for all Samples

where J is current density, E is the electric field of the insulator, Φ_b is barrier height and ξ is dielectric permittivity.

Schottky Emission:

$$J = AT^2 \exp(-q(\phi_b - \sqrt{(qE/\pi\xi)}/KT) \qquad (2)$$

where A is the effective Richardson constant.

Tunneling:

$$J \propto E^2 \exp(-(4\sqrt{2m}(q\phi_b)^{3/2})/3qhE)) \qquad (3)$$

where m is effective mass.

Observations were split into two cases for data above room temperature (300K, 350K, 400K), depending on the applied voltages. Case 1 dealt with samples in the high field region (-1.5V to -4V) while Case 2 dealt with the samples in the low field region (-0.005V to -1.5V).

Case 1

In the high field region, the I-V curve for sample 1 (Al_2O_3/Ti) fit the Schottky emission model with a barrier height of 0.59eV (Figure 4), while for sample 2 (Al_2O_3/W), the F-P emission dominated. This is logical since for W, the barrier becomes too large for the Schottky emission to dominate. A barrier height of 0.356eV was extracted from the curve fitting, as shown in Figure 5.

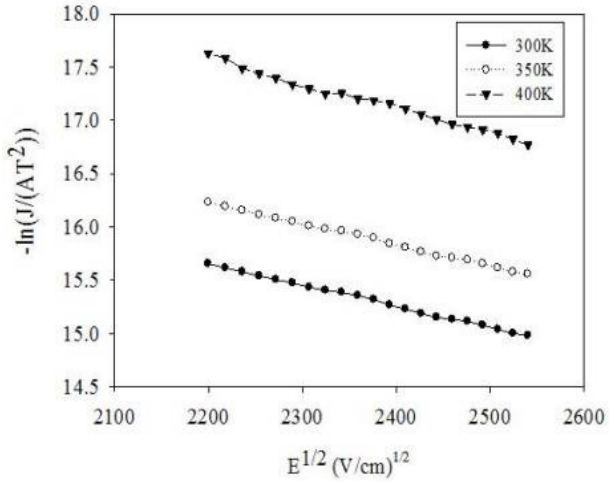


Figure 4. Schottky Emission Curve Fit for the Al_2O_3/Ti Samples at High Field

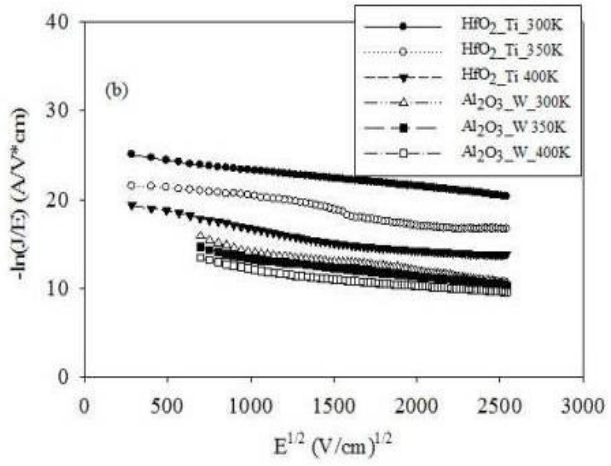


Figure 5. F-P Emission Curve Fitting for Al_2O_3/W and HfO_2/Ti at Low and High Fields

For sample 3 (HfO_2/Ti), the F-P conduction mode fit extremely well, as shown in Figure 6. The calculated Schottky barrier turned out to be greater than the extracted F-P barrier height of 0.58eV. Therefore, at high electric fields, the F-P emission seems to dominate the conduction mode. From Figure 3, it can be interpreted that the current for sample 4 (HfO_2/W) is almost constant, but high relative to samples 1 and 3. The reason could be that, in high electric fields, the I-V curve fits well both in tunneling [24] and F-P conduction mechanisms. Again, it is interesting to note from Figure 7 that the curves are not perfectly overlapping for the tunneling mechanism. This is due to the fact that the charge conduction mode is not independent. Barrier heights of 0.24eV and 0.27eV were extracted for tunneling and F-P emissions, respectively. **A s**ummary of case 1 is depicted in Table 1.

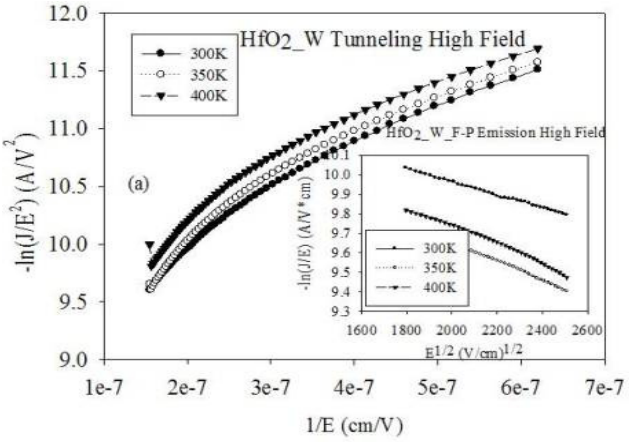


Figure 6. Curve Fitting for F-P Emission and Tunneling for HfO_2/W at High Field

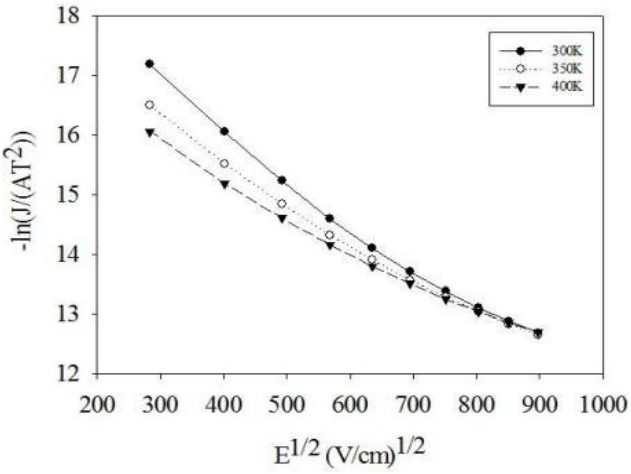


Figure 7. Schottky Emission for HfO_2/W at Low Field

Table 1. Charge Transport Mechanism and Extracted Barrier Heights of Case 1 Samples

Sample No.	*MOS Capacitors*	*High Field (-1.5V to-4V) (Case 1)*	*Barrier Height*
1	Al_2O_3/Ti	Schottky	0.59 eV
2	Al_2O_3/W	F-P	0.356 eV
3	HfO_2/Ti	F-P	0.58 eV
4	HfO_2/W	Tunneling/F-P	0.24 eV/ 0.27 eV

Case 2

In low a electric field, for sample 1, the extracted barrier height of 0.53eV for the F-P emission was less than that of the Schottky emission barrier height, as shown in Figure 8. Therefore, the F-P emission dominates in a low field for sample 1. For samples 2 and 3, the calculated Schottky barrier turned out to be greater than the extracted F-P barrier height of 0.58eV. Hence, the F-P emission dominates in a low field region, as can be seen in Figure 5. However, for sample 4, the Schottky emission acts as the dominant charge conducting mode. A barrier height of 0.49eV was extracted from the curve fitting, as shown in Figure 7. Case 2 is summarized in Table 2.

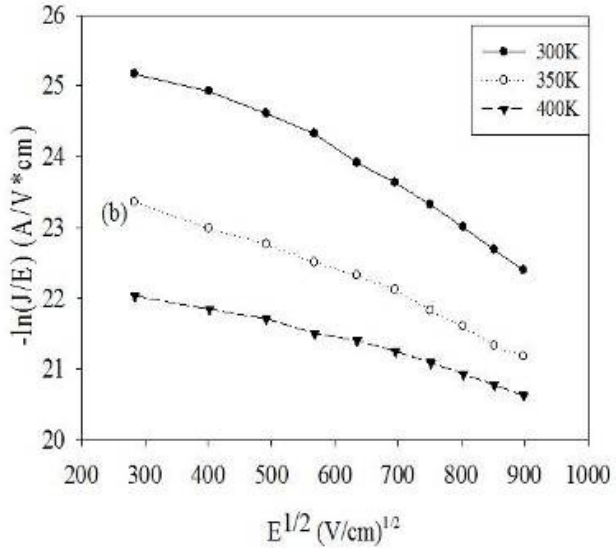


Figure 8. F-P Emission Curve Fit for the Al_2O_3/Ti Sample at Low Field

Thereafter, the modeling of the data was done for all four samples using a neural network to cover all temperature points lying in the range of 300K – 400K for voltage ranging from -4V to 4V, as previously explained in the methodology section. Figure 9 shows the actual output current and neural model output current for all the samples. It can be easily seen that the model followed the experimental set of data very closely. Table 3 shows the training error in the process of modeling the data and validation error after testing the trained model.

Table 2. Charge Transport Mechanism and Extracted Barrier Heights of Case 2 Samples

Sample No.	*MOS Capacitors*	*Low Field (Case 2) (-0.005V to-*	*Barrier Height*
1	Al_2O_3/Ti	F-P	0.53 eV
2	Al_2O_3/W	F-P	0.356 eV
3	HfO_2/Ti	F-P	0.58 eV
4	HfO_2/W	Schottky	0.49 eV

Both Figure 9 and Table 3 demonstrate excellent modeling capabilities due to a very low percentage of error. Once modeling is done, the current can be accurately calculated using this model for any given temperature range, which will be beneficial in reproducing the results without actually fabricating the device again. Hence, the model is cost effective and helps in speeding up the entire process of fabrication and testing of devices. On the other hand, the already established equations for F-P emission, tunneling etc. require the definition of a number of parameters like effective mass and barrier height, before calculating the output current. In this way, just by feeding the trained neural network with two inputs (voltage and temperature), the required output (current) can be easily established. In the future, this model would also help in the comparison of different high-k or different metal-gate-based MOS capacitors as reproducing data will be very easy.

Table 3. Training and Validating Errors Obtained from Data Modeling using Neural Network

Error	*Sample 1 (Al_2O_3/Ti)*	*Sample 2 (Al_2O_3/W)*	*Sample3 (HfO_2/Ti)*	*Sample 4 (HfO_2/W)*
Training Error	0.2309%	0.3456%	0.2104%	0.5549%
Validation Error	0.1661%	0.3045%	0.2155%	0.4671%

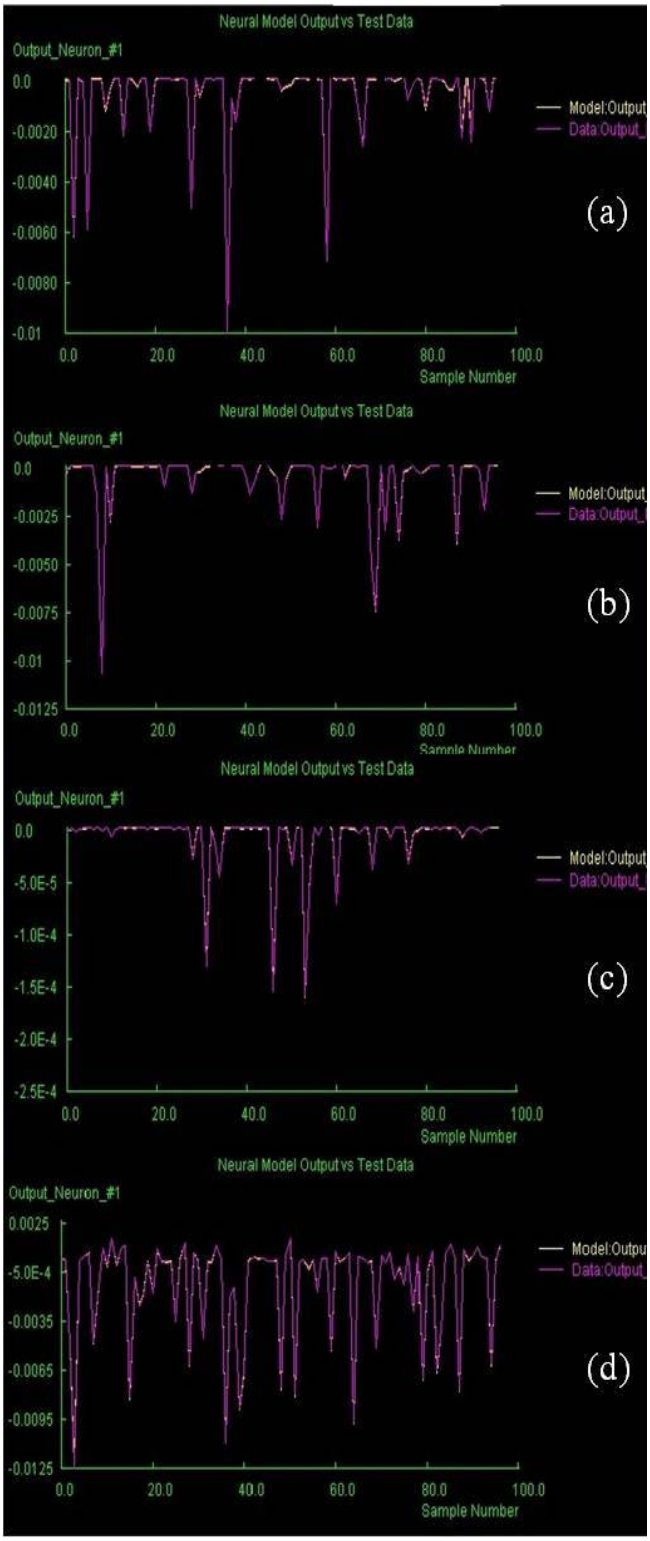


Figure 9. Sample 1 (a), Sample 2 (b), Sample 3 (c) and Sample 4 (d) Shows the Comparison of Actual and Modeled Output Current

Conclusions

The mode of conduction for the four samples consisted of a combination of F-P emission, Schottky emission and tunneling, with each mechanism dominating according to the applied bias and temperature. At low temperatures, tunneling remains as the dominant mode of conduction. However, at higher temperatures and in low field, the F-P emission is the dominant mode of conduction, except for the HfO_2/W sample. In high field, the conduction mechanism is dependent both on the electrodes and dielectrics being used. The neural-network-based modeling of the data proved to be beneficial for predicting the output (current) for different temperature points in the range being modeled without actually fabricating and testing the device. This greatly reduces the probability of manual errors due to fabrication, thereby simplifying data acquisition and improving the yield.

References

[1] Garner, C. M., Kloster, G., Atwood, G., Mosley, L. & Palanduz, A. C. (2005). Challenges for dielectric materials in future integrated circuit technologies. *Microelectronics Reliability*, 45(5-6), 919-.924

[2] Robertson, J. (2004). High dielectric constant oxides. *European Physical Journal Applied Phys*ics, 28, 265 –291.

[3] Tataroglu, A., Altındal, S. & Bulbul, M. M. (2005). Temperature and frequency dependent electrical and dielectric properties of Al/SiO2/p-Si (MOS) structure. *Microelectronic Engineering*, ,81 140-.149

[4] Chau, R., Datta, S., Doczy, M., Kavalieros, J. & Metz, M. (2003). Gate dielectric scaling for high-performance CMOS: from SiO_2 to high-k. Retrieved on October 23, 2010, from http://www .intel.com/ technology/silicon/ieee/IWGI2003.pdf

[5] Ma, T. P., He, W. & Wang, M. (2006). *Defects in High-k Gate Dielectrics*. (E. Gusev, Ed.). Springer: Netherlands.

[6] Chau, R., Datta S., Doczy, M., Doyle, B., Kavalieros, J. & Metz, M. (2004). High-κ/metal-gate stack and its MOSFET characteristics. *IEEE Electron Device Letters*, 25(6), 408-410.

[7] Jung, R. (2009). Fermi-Level Pinning at the Poly-Si/ HfO_2 Interface. *Journal of the Korean Physical Society*, 55(6), 2501-.2504

[8] Gusev, E. P., Buchanan, D. A., Cartier, E., Kumar, A., DiMaria, D., Zafar, et al. (2001). *Ultrathin high-K gate stacks for advanced CMOS devices*. IEDM Technical Digest.

[9] Xiong, K., Peacock, P. W. & Robertson, J. (2005). Fermi level pinning and Hf–Si bonds at HfO_2: Poly-

crystalline silicon gate electrode interfaces. *Applied Physics Leters,*86 .(1) .

[10] Pantisano, L., Afanas'ev, V. V., Cimino, S., Adelmann, C., Goux, L., Chen, Y. Y., et al. (2011). Towards barrier height modulation in HfO2/TiN by oxygen scavenging – Dielectric defects or metal induced gap states? *Microelectronic Engineering,*88 , 1251-.1254

[11] Fischetti, M. V., Neumayer, D. A. & Cartier, E. A. (2001). Effective electron mobility in Si inversion layers in metal–oxide–semiconductor systems with a high-κ insulator: The role of remote phonon scattering**.** *Journal of Applied Physics*,90 (9), 4587–4608.

[12] Tan, S. Y. (2007). Challenges and performance limitations of high-k and oxynitride gate dielectrics for 90/65 nm CMOS technology. *Microelectronic Engineeering,* 38, 783-786.

[13] Fu, C. H., Chang, K. S., Chang, Y. A., Hsu, Y. Y., Tzeng, T. H., Wang, T. K.. et al. (2011). A low gate leakage current and small equivalent oxide thickness MOSFET with Ti/HfO_2 high-k dielectric. *Microelectronic Engineering*,88 (7). 1309-.1311

[14] Luo, Z. J., Guo, X. & Ma, T. P. (2001). Temperature dependence of gate currents in thin Ta_2O_5 and TiO_2 films. *Applied Physics Letters*,79 , 2803-.2804

[15] Zhu, W. J., Ma, T. P., Tamagawa, T., Kim, J. & Di, Y. (2002). Current transport in metal/Hafnium oxide Silicon structure. *IEEE Electron Device Letters*,23 , 97-.99

[16] Lin, C., Kang, J., Han, D., Tian, D., Wang, W., Zhang, J. et al. (2003). Electrical properties of Al_2O_3 gate dielectrics. *Microelectronic Engineering*,66 , 830-.834

[17] Negarestani, A., Setayeshi, S., Maragheh, M. G. & Akashe, B. (2003). Estimation of the radon concentration in soil related to the environmental parameters by a modified Adaline neural network. *Applied Radiation and Isotopes*,,(2)58 269-.273

[18] Kweon, K. E., Lee, J. H., Ko, Y, Jeong, M., Myoung, J. & Yun, I. .(2007) Neural network based modeling of HfO_2 thin film characteristics using Latin Hypercube Sampling. *Science Direct*, 32(2), 358–363.

[19] Ojala, P., Saarinen, J., Elo, P. & Kaski, K. (1995). Novel technology independent neural network approach on device modeling interface. *IEE Proceedings, Circuits Devices and Systems,*142 (1), 74-82.

[20] Creech, G. L. & Zurada, J. M. (1999). *Neural network modeling of GaAs IC material and MESFET device characteristics*. John Wiley and Sons, Inc.

[21] Sze, S. & Ng, K. (2007). *Physics of Semiconductor Devices*. (3rd ed.). New Jersey: Wiley.

[22] Schroder, D. (2006). *Semiconductor Material and Device Characterization*. (3rd ed.). New Jersey: Wiley.

[23] Ranuarez, J. C., Deen, M. J. & Chen, C. (2006). A review of gate tunneling current in MOS devices. *Microelectronics Reliability*, 46(12), 1939-.1956

Biographies

MADHUMITA CHOWDHURY is pursuing her M.S. degree in EECS department at the University of Toledo. Her research interests are fabrication and characterization of resistive RAMS (memristors), compound semiconductors and high-k material-based devices. She is also interested in fabrication of flexible/eco-friendly materials like Graphene to be used as electrodes. Madhumita Chowdhury can be reached at madhumita.chowdhury@rockets.utoledo.edu

BRANDEN LONG obtained Bachelor of Science in Electrical Engineering from the University of Michigan, in 2007, and M.S. in Electrical Engineering at the University of Toledo in 2009. He is currently working towards his Ph.D. degree in the EECS Department at the University of Toledo. His research interests include fabrication and characterization of advanced memory and logic devices, compound semiconductors, thin film transistors, and exploring novel switching mechanisms in semiconductor devices for beyond CMOS applications. Branden Long can be reached at branden.long@rockets.utoledo.edu

RASHMI JHA is an assistant professor in the department of Electrical Engineering and Computer Science. Her research expertise lies in the areas of high permittivity (high-k) transition metal oxide and metal gate electrodes based advanced Complementary Metal Oxide Semiconductor Field Effect Transistor (CMOS FET) fabrication and characterization for the next-generation analog, digital, and memory applications. Dr. Jha can be reached at Rashmi.jha@utoledo.edu

VIJAY DEVABHAKTUNI (S'97, M'03, and SM'09) is an Associate Professor in the Electrical Engineering and Computer Science Department, University of Toledo, Toledo, Ohio, USA. He received the B.Eng. degree in electrical and electronics engineering and the M.Sc. degree in physics both from the Birla Institute of Technology and Science, Pilani, Rajasthan, India, in 1996, and the Ph.D. degree in electronics from the Carleton University, Ottawa, Ontario, Canada, in 2003. His research interests include computer aided design (CAD), image/signal processing, modeling, RF/microwave design, optimization and wireless sensing. He can be reached at Vijay.Devabhaktuni@utoledo.edu

A NUMERICAL METHOD FOR PERMEABILITY ESTIMATION

Dacun Li, University of Texas of the Permian Basin

Abstract

Correlation between porosity and permeability for a certain rock type is a basic procedure used in core-data interpretation. However, the correlation may not always be satisfactory due to pore heterogeneity and pore geometry. In a reservoir, it is very common for rocks to have similar porosities but different permeabilities. Apparent formation factor was defined as true resistivity divided by water resistivity. In previous work, curves of the apparent formation factor versus water saturation were used to interpolate permeabilities. Unfortunately, the accuracy was very poor at high water saturation because the curves were horizontal in that region. Even in the regions which are not horizontal, chart reading can be very subjective and accuracy is limited. In this study, a numerical method was proposed in order to solve this problem. Based on combining a correlation between the formation resistivity factor and permeability proposed by Ogbe and Bassiouni [1], Archie's Equation, and the apparent formation factor definition, a new equation was derived in order to estimate permeabilities for clean formations from resistivity logs.

An algorithm was developed to carry out the procedures of the estimation, an example of which is included here in order to illustrate how to use the algorithm. This example shows that the permeability estimate from this study was more accurate than that from the previous work. In addition, this study yielded an objective permeability estimate while, the previous work gave only a subjective estimate. A computer program developed from the algorithm can be incorporated into a reservoir simulator as an improved way to input permeability values.

Introduction

Permeability is one of the most important properties to estimate for a reservoir [2]. A good correlation between the porosity and permeability of a certain rock type is desirable for reservoir characterization or simulation. Unfortunately, the correlation may not always be satisfactory due to pore heterogeneity and pore geometry [3]. It is very common for rocks in a reservoir to have similar porosities but different permeabilities.

By analyzing laboratory measurements on 155 sandstone samples from three different oil fields in North America, Timur [4] found the empirical correlation

$$k = f^{4.4}/(S_{wr})^2 \quad (1)$$

where k is permeability, f is porosity, and S_{wr} is residual water saturation. Morris and Biggs [5] and Coates and Dumanoir [6] also developed correlations between k, f and S_{wr}. However, these methods are only applicable at irreducible water saturation. Saner et al. [7] defined the apparent formation resistivity factor as true formation resistivity divided by water resistivity. Saner et al. [3] used curves of the apparent formation resistivity factor versus water saturation to interpolate permeabilities. However, it was difficult to get an accurate permeability reading from the plots, especially at high water saturations, because the curves became asymptotic or horizontal. This current study extends the work by Saner et al. [3] by developing a numerical method to solve the accuracy problem of the plot reading. A new equation was derived and an algorithm developed in order to estimate permeabilities for clean formations from resistivity logs.

Approach

In the following sections, the derivation of the equation and the development of the algorithm are presented. An example is also provided to illustrate the application of the algorithm.

Equation and Algorithm Development

By Combining

$$\tau = F^b \quad (2)$$

where F is the formation resistivity factor, τ is tortuosity, and b is an exponent depending on the rock texture, and

$$k \propto \frac{1}{\tau} \quad (3)$$

where k is permeability, a correlation between the formation resistivity factor, F , and permeability, k, was found such that

$$F = Ak^{-B} \quad (4)$$

where A and B are constants for a specific formation.

According to Archie's Law,

$$F = S_w^{\,n} \frac{R_t}{R_w} \tag{5}$$

where S_w is water saturation, R_t is true resistivity, n is the water saturation exponent, and R_w is water resistivity. Let R_o refer to formation resistivity when water saturation is 100%. At water zones $S_w = 1$, R_o is equal to R_t. From Equation (5), it can be seen that only when $S_w = 1$ will $F = R_t / R_w$. Apparent formation factor, F_a, was defined by Saner et al. [7] as

$$F_a = \frac{R_t}{R_w} \tag{6}$$

From Equations (4) - (6), one gets

$$S_w^{\,n} F_a = Ak^{-B} \tag{7}$$

Taking a logarithm for both sides of Equation (7), and then rearranging, yields

$$\log(F_a) = -n\log(S_w) + (\log(A) - B\log(k)) \tag{8}$$

It can be inferred from Equation (8) that $\log(F_a)$ is linear to $\log(S_w)$ for a certain group of core samples, which have constant n, A, B and k values. If many core samples are obtained, they can be grouped according to the procedures similar to those provided by Saner et al. [3], and a series of straight lines of $\log(F_a)$ versus $\log(S_w)$ can be plotted. The difference of intercepts between different lines is due to the difference of the permeabilities of the groups represented by the lines. The higher the permeability, the lower the $\log(F_a)$-axis intercept of the line. Based on the above discussion, it can be inferred that the curves of $\log(F_a)$ versus $\log(S_w)$, plotted from different groups of core samples, can be used to estimate permeability. Also the more data obtained, the more accurate the final estimate.

For a certain part of the formation, true resistivity can be obtained from a deep induction log, and water resistivity can be obtained from a water catalog, water analysis, SSP (static spontaneous potential), or other kinds of cross-plots. Then, the apparent formation resistivity factor can be calculated from Equation (6). Porosity, f , can be obtained from the density log, sonic log or neutron log. If parameters such as cementation factor *a*, cementation exponent *m* and saturation exponent *n*, are known, then for a clean formation (without shale), water saturation can be calculated from Archie's Equation as follows:

$$S_w = \left[\frac{aR_w}{\phi^m R_t} \right]^{\frac{1}{n}} \tag{9}$$

McCoy and Grieves [8] present procedures to calculate water saturation at Prudhoe Bay. If parameters such as *a*, *m* and *n* are unknown for a formation, procedures similar to those illustrated by McCoy and Grieves [8] can be applied to solve for these parameters. Alfosail and Alkaabi [9] developed an equation to calculate water saturation in shaly formations. If the formation is a shaly formation, it is required that a suitable equation be developed to calculate the water saturation following similar steps [9].

Suppose point $\log(S_w)$, $\log(F_a)$ is located between the two straight lines defined by the following two equations:

$$\log(F_{a1}) = -n_1 \log(S_{w1}) + (\log(A) - B\log(k_1)) \tag{10}$$

and

$$\log(F_{a2}) = -n_2 \log(S_{w2}) + (\log(A) - B\log(k_2)) \tag{11}$$

where k_1 and k_2 refer to the permeabilities of the individual straight lines, and n_1 and n_2 are the slopes. In Appendix A, a new equation is derived to estimate permeabilities for clean formations from resistivity logs, that is,

$$k = k_1 \left[\frac{k_2}{k_1} \right]^{\frac{b_1 - b}{b_1 - b_2}} \tag{12}$$

where $b_1 = \log(A) - B\log(k_1)$, $b_2 = \log(A) - B\log(k_2)$. If $n_1 \neq n_2$, then b is calculated by

$$b = \frac{x_2 y_1 - x_1 y_2}{x_2 - x_1} \tag{13}$$

where $x_2 = \log(S_w)$, $y_2 = \log(F_a)$, and

$$x_1 = \frac{b_2 - b_1}{n_2 - n_1} \tag{14}$$

$$y_1 = \frac{n_2 b_1 - n_1 b_2}{n_2 - n_1} \tag{15}$$

If $n_1 = n_2$, Equations (14) and (15) are not valid. If $n_1 = n_2 = n'$, then b is calculated by

$$b = n' \log(S_w) + \log(F_a) \tag{16}$$

Based on the above analysis, an algorithm was developed to estimate permeability, as shown in Figure 1. In the following two sections, (Synthetic Example and Analysis), an example is given to illustrate the permeability estimation process for a carbonate formation.

Synthetic Example

Suppose that measurements for six groups of core samples are available from a carbonate formation and each

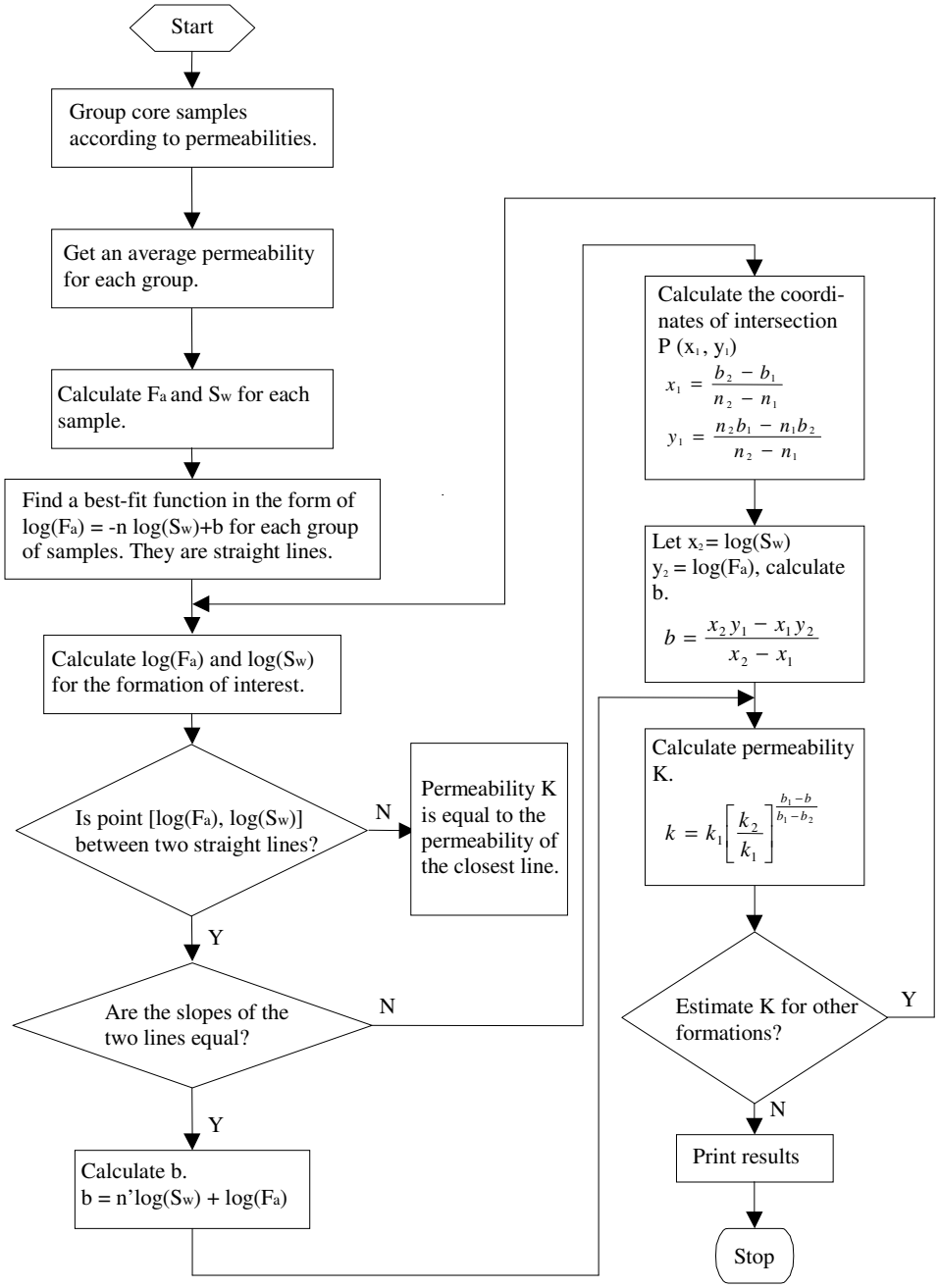


Figure 1. Algorithm to Estimate Permeabilities for Clean Formations from the Resistivity

group has a different average permeability, K_{ave}. Correlations of the apparent formation resistivity factor F_a versus water saturation, S_w , are listed in Table 1, where data are created by following the example given by Saner et al. [3]. There are log measurements for a part of the formation, and then water saturation and apparent formation resistivity factor values that can be calculated from Equations (9) and (6), respectively. Suppose further that the water saturation is found to be 0.50 and the apparent formation resistivity factor is 100.00, then estimate the permeability for this part of the formation.

Analysis

Following the above procedures, the permeability can now be estimated. Curves of apparent formation resistivity

Table 1. F_a vs S_w for Six Groups of Core Samples

Sw	Fa					
	Group 1 K_{ave}=0.01 md	Group 2 k_{ave}=0.15 md	Group 3 k_{ave}=1.5 md	Group 4 K_{ave}=10 md	Group 5 K_{ave}=85 md	Group 6 K_{ave}=750 md
0.07					1681.77	1317.99
0.08					1326.37	1042.14
0.09					1075.99	847.17
0.10					892.54	703.88
0.11	82037.79	31771.31	4962.92	1770.31	897.84	795.26
0.12	63578.29	20724.38	3748.99	766.72	410.26	360.80
0.13	50094.78	18290.74	5174.57	908.39	335.01	573.73
0.14	41443.91	18960.95	1025.63	1082.82	615.24	260.51
0.15	33484.65	9619.26	3391.17	301.58	627.89	465.01
0.16	28659.62	12725.17	2114.59	887.76	269.11	207.99
0.17	23324.28	7137.20	2188.68	280.15	343.82	366.84
0.18	21575.76	7568.38	1927.41	252.68	310.47	155.37
0.19	16588.56	7736.30	1709.04	663.06	281.90	314.66
0.20	17359.51	3342.57	1234.79	184.57	198.24	130.02
0.22	10657.45	5192.25	1693.54	315.95	292.99	255.92
0.23	9498.35	2993.49	817.42	347.98	160.44	113.69
0.24	10962.10	3469.12	1016.51	236.32	185.77	190.96
0.21	12328.30	3982.76	967.99	346.88	235.78	150.92
0.22	9769.02	5892.25	1433.54	354.95	266.99	210.92
0.23	11103.78	2893.49	838.42	288.98	133.44	129.69
0.25	7136.82	3905.60	628.29	244.44	221.72	109.50
0.27	7495.67	2520.66	782.25	269.45	219.55	122.72
0.28	4896.33	2283.95	721.47	116.70	111.08	115.11
0.29	6195.73	1676.64	667.31	230.46	104.52	108.22
0.30	5566.46	2194.25	418.85	119.52	180.74	101.96
0.32	3157.54	1590.14	734.10	187.92	81.16	72.02
0.33	3006.30	1762.84	300.64	140.00	127.22	101.23
0.35	3314.17	891.11	539.23	167.40	64.73	64.75
0.36	2938.82	879.39	313.56	117.57	126.08	97.99
0.38	2541.43	1277.84	444.81	138.47	52.80	38.28
0.39	2415.72	620.97	266.28	61.11	78.09	87.28
0.41	2788.32	522.04	388.94	90.55	71.42	36.87
0.42	711.35	949.67	223.82	135.28	68.41	86.42
0.44	1937.33	670.52	341.03	49.58	82.96	32.99
0.45	1383.01	860.88	251.16	75.12	39.49	68.98
0.47	1201.82	460.71	228.01	98.84	88.97	26.30
0.48	1464.94	529.60	217.58	46.99	53.90	44.61
0.50	919.47	774.10	148.70	63.80	25.12	41.52
0.51	1133.57	249.32	220.14	58.43	48.37	30.10
0.52	1151.93	526.27	182.10	36.19	65.73	68.76
0.53	995.48	315.81	174.55	73.09	31.16	14.48
0.54	966.84	484.80	217.44	52.09	55.68	46.27
0.56	963.76	302.67	154.43	48.43	30.94	16.02
0.58	891.45	367.02	94.84	27.13	38.45	31.98
0.60	721.10	248.18	132.46	42.16	36.19	40.13
0.62	578.20	322.58	123.15	55.48	47.14	13.44
0.64	546.53	175.75	154.75	37.04	17.25	36.90
0.66	651.07	223.32	74.16	53.82	48.53	14.48
0.68	438.97	328.95	129.28	32.80	28.95	37.18
0.70	385.29	78.38	94.02	20.94	36.49	15.98
0.72	491.11	210.38	58.31	29.24	26.14	29.87
0.74	396.12	163.75	83.09	27.68	17.89	8.84
0.76	473.46	202.33	100.30	33.23	23.73	26.88
0.78	227.41	141.97	93.91	12.90	19.66	19.00
0.80	326.44	94.55	40.86	23.67	31.66	12.17
0.82	329.27	95.97	84.13	12.52	20.72	24.40
0.84	278.77	146.12	46.68	21.46	9.85	13.67
0.86	226.74	37.95	79.48	25.47	24.03	16.00
0.88	258.56	90.36	37.52	19.55	8.27	15.37
0.90	233.27	124.31	53.76	18.68	17.55	19.77
0.92	169.43	64.74	32.20	11.88	25.87	11.21
0.94	206.68	85.60	48.80	20.12	16.24	18.68
0.96	224.12	115.85	69.57	12.41	20.64	13.18
0.98	147.16	57.45	44.48	20.75	8.08	8.72
1.00	195.98	89.38	15.53	12.12	20.54	12.27

factor versus water saturation are plotted in Figure 2, along with the regression fit equations. Figure 3 shows the curves of $\log(F_a)$ versus $\log(S_w)$ plotted from the functions obtained for the six groups of core samples. From Figure 2, it can be determined that point (0.50, 100) is between the curve for the average permeability of 1.5md and that the average permeability of 10md. That is, in Figure 3, point log(0.5), log (100) is between the straight lines defined by the following two equations:

$$\log(F_{a1}) = -2.1866\log(S_{w1}) + 1.6096 \tag{17}$$

and

$$\log(F_{a2}) = -1.998\log(S_{w2}) + 1.162 \tag{18}$$

From Equations (14) and (15), the coordinates of the intersection for the two straight lines was calculated to be (2.3732, -3.5798). The measurement is at point log(0.50), log(100.00), that is, point (-0.3010, 2.0000). The intercept for the straight line which passes points (2.3732, -3.5798) and (-0.3010, 2.0000) can be calculated by Equation (13). The intercept (b) is 1.3720. With k_1 = 1.5md, k_2 =10md, b =1.3720, b_1 =1.6096 and b_2 = 1.162, the permeability for this part of the formation, k, was calculated from Equation (11) and found to be 4.106md.

Results and Discussions

From the example just presented, the authors feel that the numerical method does improve the accuracy of the permeability estimation. From Figure 2, to estimate the permeability by reading the plot, 3md can be used as an estimate, albeit a very subjective one. Using the numerical method, the permeability was estimated to be 4.106md. Figure 4 is a plot of the apparent formation factor versus water saturation from the previous work [3]. With a water saturation of 0.5 and apparent formation factor of 100, from Figure 4, one estimate of permeability may give 5md while another may give 3md. The results are very subjective and inaccurate. The accuracy of the estimate from the work presented here is much improved over the previous work and improves

permeability estimates by several digits beyond the decimal point. The importance of this work is to improve the accuracy of permeability estimation and reduce human errors in figure reading. In addition, the new equation and algorithm developed in the permeability estimation process can be programmed into computers, thus speeding up the estimation process.

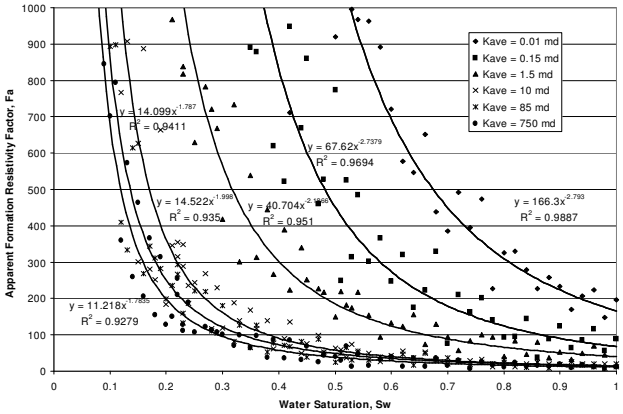


Figure 2. Plot of Apparent Formation Resistivity F_a versus Water Saturation S_w

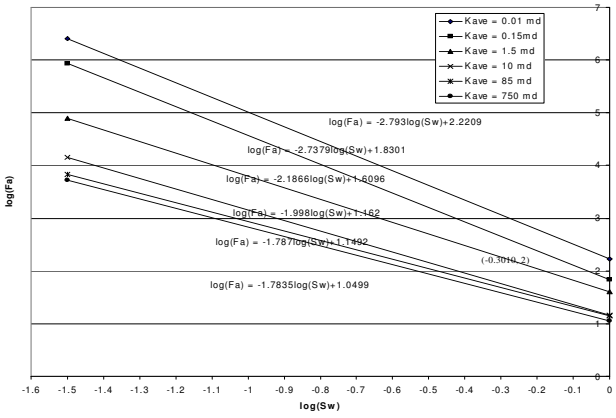


Figure 3. Log(F_a) vs Log(S_w)

As mentioned previously in reference to Equation (4), A and B are constants for a certain formation. Now let us investigate what parameters A and B may be related to. The Carman-Kozeny equation is

$$k = \frac{\phi}{(k_z \tau) S_{pv}^{\ 2}} \tag{19}$$

where k_z is Kozeny's constant and S_{pv} is the internal surface area of the pores per unit of pore volume. The generalized τ-F relationship is in the form of

$$\tau = (F\phi)^y \tag{20}$$

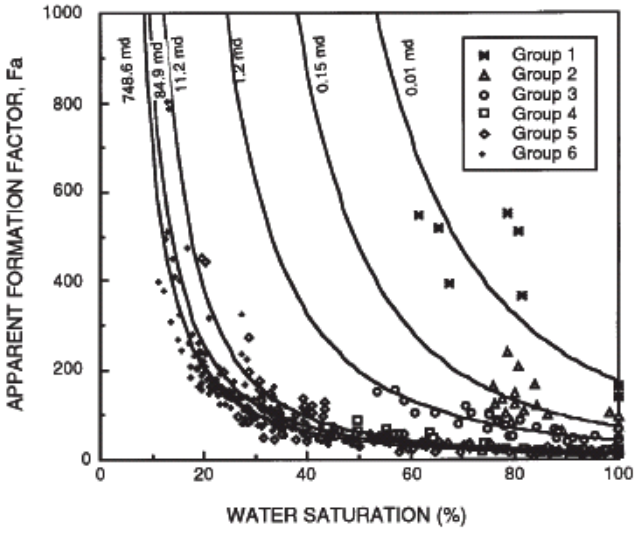


Figure 4. Apparent Formation Factor vs Water Saturation for Various Permeability Groups [3]

where y is an exponent. Combining Equations (18) and (19) gives

$$k = \frac{\phi^{1-y}}{k_z F^y S_{pv}^2} \tag{21}$$

Substituting the Salem & Chilingarian [10] relationship

$$k_z = 2.24(F\phi) \tag{22}$$

into Equation (21) leads to

$$k = \frac{\phi^{-y}}{2.24 F^{y+1} S_{pv}^2} \tag{23}$$

Rearranging Equation (23) yields

$$F = \left[\frac{\phi^{-y}}{2.24 S_{pv}^2}\right]^{\frac{1}{1+y}} k^{-\left[\frac{1}{1+y}\right]} \tag{24}$$

Comparing Equation (24) with Equation (4), we have

$$B = -\frac{1}{1+y} \tag{25}$$

and

$$A = \left[\frac{\phi^{-y}}{2.24 S_{pv}^2}\right]^B \tag{26}$$

Equations (25) and (26) show that the constant, B , is related only to y, the exponent in the generalized τ-F relationship, while the other constant, A , is a function of y, the pore

-volume-based specific surface, S_{pv}, and porosity f. For a certain formation, y, S_{pv} and f are referred to average properties and can be considered constants; thus, from Equations (25) and (26), it can be said that A and B are constants.

Conclusions

From this study, three conclusions can be reached:

1. A new equation was derived and an algorithm developed to calculate permeabilities for clean formations from resistivity log measurements. By applying the algorithm, accuracy was improved in the estimation of permeability.
2. The algorithm can be incorporated into a reservoir simulator so that it will provide an improved permeability input.
3. Equations to predict the constants in the correlation of the formation resistivity factor versus permeability were derived. Parameters related to the constants were also found.

Nomenclature

English Symbols

A = constant

a = cementation factor

b = exponent of formation resistivity factor; intercept

B = constant

F = formation resistivity factor

k = permeability, L^2, md; Kozeny's constant

K_{ave} = average permeability, L^2, md

m = cementation exponent

n = water saturation exponent

R = resistivity, W.m

S = saturation; specific surface, 1/L, 1/m

x = $\log(S_w)$ axis

y = $\log(F_a)$ axis; exponent

Greek Symbols

Δ = intercept difference

ϕ = porosity

τ = tortuosity

Subscripts

1 = property of line 1

2 = property of line 2

a = apparent

o = 100% water saturation

pv = pore volume

r = residual

t = True

w = water

z = Kozeny

References

[1] Ogbe, D. & Bassiouni, Z. (1978). Estimation of Aquifer Permeabilities from Electrical Well Logs. *The Log Analyst*, 19(5), 21-27.

[2] Lopez, C. & Davis, T. L. (2011). *Permeability Prediction and its Impact in Reservoir Modeling. Postle Field, Oklahoma*. AAPG Search and Discovery Article #90129, presented at the AAPG Southwest Section Meeting, Ruidoso, New Mexico.

[3] Saner, S., Kissami, M. & Nufaili, S. A. (1997). Estimation of Permeability from Well Logs Using Resistivity and Saturation Data. *SPE Formation Evaluation*, 12(1). 27-31.

[4] Timur, A. (1968). An Investigation of Permeability, Porosity, and Residual Water Saturation Relationship for Sandstone Reservoirs. *The Log Analyst*, 9(4).

[5] Morris, R. L. & Biggs, W. P. (1967). Using Log-Derived Values of Water Saturation and Porosity. *Proceedings of the SPWLA Eight Annual Logging Symposium*.

[6] Coates, G. R. & Dumanoir, J. L. (1973). A New Approach to Improved Log-Derived Permeability. *Proceedings of the SPWLA Fourteen Annual Logging Symposium*.

[7] Saner, S., Orcan, A. & Wajid, M. (1994). Apparent Cementation Factor Concept for Water Saturation Determination from Well Logs. *The Arabian Journal for Science and Engineering*, 19(3).

[8] McCoy, D. D. & Grieves, W. A. (1997). Use of Resistivity Logs to calculate Water Saturation at Prudhole Bay. *SPE Reservoir Engineering*, 12(1), 45-51.

[9] Alfosail, K. A. & Alkaabi, A. U. (1997, March). *Water Saturation in Shaly Formation*. Paper presented at the Middle East Oil Show and Conference, Bahrain.

[10] Chilingarian, G., Torabzadeh, J., Rieke, H., Metghal-

chi, M., & Mazzullo, S. (1992). Interrelationships Among Surface Area, Permeability, Porosity, Pore Size, and Residual Water Saturation. In Chilingarian, G., Mazzullo, S., & Rieke, H. (Eds.), *Carbonate Reservoir Characterization: A Geologic-Engineering Analysis* (pp. 379-397). Amsterdam: Elsevier.

Appendices

Derivation of the New Equation

Suppose point $\log(S_w)$, $\log(F_a)$ is located between the two straight lines defined by the following two equations:

$$\log(F_{a1}) = -n_1 \log(S_{w1}) + (\log(A) - B\log(k_1)) \quad \text{(A-1)}$$

and

$$\log(F_{a2}) = -n_2 \log(S_{w2}) + (\log(A) - B\log(k_2)) \quad \text{(A-2)}$$

The difference of the intercepts for these two straight lines is

$$\Delta = B(\log(k_2) - \log(k_1)) \quad \text{(A-3)}$$

Now, if we let

$$b_1 = \log(A) - B\log(k_1) \quad \text{(A-4)}$$

and

$$b_2 = \log(A) - B\log(k_2) \quad \text{(A-5)}$$

then from Equations (A-3), (A-4) and (A-5) we get

$$B = \frac{b_1 - b_2}{\log(k_2) - \log(k_1)} \quad \text{(A-6)}$$

If in Equations (A-1) and (A-2) $n_1 \neq n_2$, then the lines are not parallel and the two straight lines defined by Equations (A-1) and (A-2) cross at point P. From Equations (A-1), (A-2), (A-4) and (A-5), the coordinates of point P (x_1, y_1) can be calculated as

$$x_1 = \frac{b_2 - b_1}{n_2 - n_1} \quad \text{(A-7)}$$

$$y_1 = \frac{n_2 b_1 - n_1 b_2}{n_2 - n_1} \quad \text{(A-8)}$$

Point $\log(S_w)$, $\log(F_a)$ is located between the two straight lines defined by the Equations (A-1) and (A-2). Next, let

$$x_2 = \log(S_w) \quad \text{(A-9)}$$

and

$$y_2 = \log(F_a) \quad \text{(A-10)}$$

then the straight line which passes points (x_1, y_1) and (x_2, y_2) has a function of

$$y = \frac{y_2 - y_1}{x_2 - x_1}(x - x_1) + y_1 \quad \text{(A-11)}$$

Its y intercept is

$$b = \frac{x_2 y_1 - x_1 y_2}{x_2 - x_1} \quad \text{(A-12)}$$

The difference between the intercept of the straight line defined by Equation (A-11) and that of the line defined by Equation (A-1) is

$$b_1 - b = -B[\log(k_1) - \log(k)] \quad \text{(A-13)}$$

From Equations (A-6) and (A-13), k can be solved as

$$k = k_1 \left[\frac{k_2}{k_1}\right]^{\frac{b_1 - b}{b_1 - b_2}} \quad \text{(A-14)}$$

From Equation (A-14), when $b = b_1$, it can be seen that $k = k_1$; and when $b = b_2$, $k = k_2$. So, the calculation of permeability is effective for the whole interval $b \in [b_2, b_1]$. If $n_1 = n_2 = n$, then the two straight lines defined by Equations (A-1) and (A-2) are parallel to each other. From point (x_2, y_2), a straight line can be drawn parallel to these two lines. Letting b be the intercept of this new line, then

$$\log(F_{a1}) - \log(F_a) = b_1 - b = -B[\log(k_1) - \log(k)] \quad \text{(A-15)}$$

and

$$\log(F_{a1}) - \log(F_{a2}) = b_1 - b_2 = -B[\log(k_1) - \log(k_2)] \quad \text{(A-16)}$$

From Equations (A-15) and (A-16), k can be solved as

$$k = k_1 \left[\frac{k_2}{k_1}\right]^{\frac{b_1 - b}{b_1 - b_2}} = k_1 \left[\frac{k_2}{k_1}\right]^{\frac{\log(F_{a1}) - \log(F_a)}{\log(F_{a1}) - \log(F_{a2})}} \quad \text{(A-17)}$$

where $b = n\log(S_w) + \log(F_a)$; b_1 and b_2 were defined in Equations (A-4) and (A-5); $\log(F_{a1})$ was calculated from Equation (A-1), assuming that $S_{w1} = S_w$, $\log(F_{a2})$ was calculated from Equation (A-2) and that $S_{w2} = S_w$.

Furthermore, from Equation (A-17), when $b = b_1$, it can be seen that $k = k_1$; and when $b = b_2$, $k = k_2$. So, the calculation of permeability is also effective for the whole interval $b \in [b_2, b_1]$. The first part of Equation (A-17) is the same as Equation (A-14). If the permeability is calculated from in-

tercept differences, for both cases, $n_1 = n_2$ and $n_1 \neq n_2$, the same equation can be used. If n_1 and n_2 are equal to each other, another method for calculating the permeability would be to use the second part of Equation (A-17).

SI Metric Conversion Factors

md ´ 9.869 233 E-04 = mm^2

Biography

DACUN LI is featured in *Who's Who in America 2011 (65th Edition)*. He is Coordinator and Assistant Professor of Petroleum Engineering Program at the University of Texas of the Permian Basin, and Editorial Review Committee member of the Society of Petroleum Engineers (SPE). Holding three degrees respectively in three different areas (a Bachelor's degree in Aerospace Engineering, a Master's degree in Health Physics, and a Ph.D. in Petroleum Engineering), he has international, academic, and industrial work experiences. He was one of the main characters in the TV documentary titled *Red Capitalism* (1994 Canada's Golden Sheaf Award winner), produced by Canadian Broadcasting Corporation (CBC) in July 1993. With a personable character, Dr. Li likes singing, dancing, photographing, practicing calligraphy, sledding, jogging, playing table tennis, and traveling. He can be reached at li_d@utpb.edu

DESIGN AND ANALYSIS OF ULTRASONIC NDT INSTRUMENTATION THROUGH SYSTEM MODELING

T. Parthipan, Brunel University; P. Mudge, TWI Ltd.; R. Nilavalan, Brunel University; W. Balachandran, Brunel University

Abstract

System modeling techniques were used to perform a power analysis of a battery-operated non-destructive testing system to reliably predict the required power source for the system. Topology-specific and equivalent circuit models written in SPICE were employed for modeling this multi-engineering discipline system. Crucial system constraints such as power consumption and power dissipation were simulated and evaluated. Presented here is the modeling work carried out on the electromechanical load, excitation-circuits and high-voltage power supply. Simulation schematics were translated into hardware schematics and prototyped. Measurements on prototyped hardware are also presented for comparison with simulation results and model evaluation. Close mapping of simulation results to real hardware was obtained for topology-specific models; agreement within 20% was achieved for equivalent circuit models. The system model developed in this study is currently being used as a virtual test platform for verifying various design methodologies and foreseeing uncertainties.

Introduction

Structures which are installed and operate in harsh environments in extreme conditions are likely to fail prematurely. Reliable structural monitoring instrumentation is crucial for monitoring the condition of these structures. Long Range Ultrasonic Testing (LRUT) is a novel non-destructive testing (NDT) method used in the detection of volumetric defects such as gross corrosion [1].

The LRUT method uses elastic waves (sound waves) in the kilohertz range and allows couplant-free transducer coupling to the test specimen. In LRUT, a pulse-echo method is often used so that access to a single location is enough to inspect long range. The A-Scan (top) and its corresponding A-Map of a pipeline being inspected on both sides from a single test location are shown in Figure 1. A-scans are plots of signal amplitude against time and A-Map is the plan view of the pipe length. In this method, a transducer or an array of transducers excites sound waves, which propagate into the material of the test specimen. The propagation is constrained by the upper and lower surfaces of the specimen, hence the term guided wave. These interact with features such as defects and reflect back to the same transducer that captures the echo signal [2].

To date, the most common application of LRUT is in the in-situ inspection of industrial pipelines [1], [2]. The effectiveness and the economic viability of the LRUT method led industry to broaden its applications to include condition monitoring of large remote structures or those with limited access for maintenance such as offshore wind-farm turbine towers, tanks and floating production storage and off-loading vessels (FPSOs). These applications require remotely installed, distributed-sensor networks based on the LRUT technique.

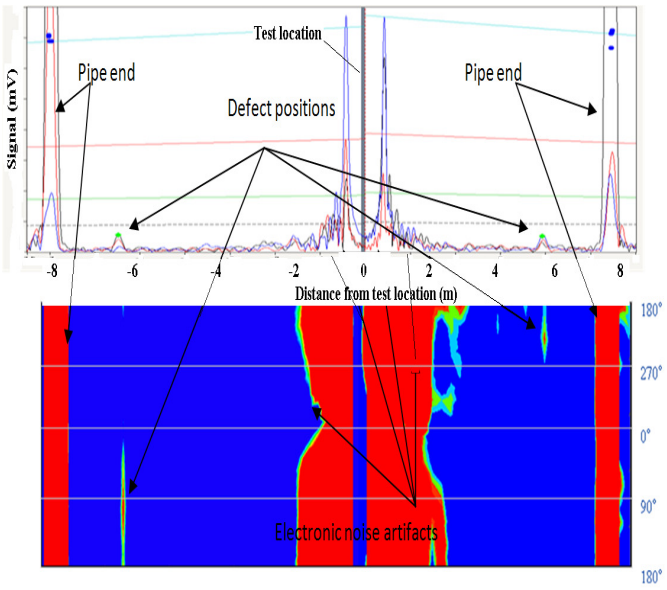


Figure 1. A-Scan and A-Map of the LRUT Applied Pipe [2]

A pulser receiver unit (PRU) allows the LRUT technique to be implemented in remote locations. It is an electronic instrument powered by rechargeable lithium ion (Li-Ion) batteries. Uninterrupted structural monitoring relies on reliable instrumentation. There is evidence that premature failure of remote sensing instrumentation is mainly caused by inadequate power sources [3], [4]. Reliably estimating the power budget of the system is, therefore, crucial for allocating adequate power sources for LRUT instrumentation.

Power consumption can vary for different functions and scenarios. Power analysis on LRUT hardware for various scenarios and functionalities can enable one to specify an adequate power source for the instrumentation that would last for the scheduled inspection period. However, practical experimentation using hardware is not always possible at early stages of the project, nor is it feasible to run such experiments for all scenarios. Computer system modeling of the LRUT system forms a virtual test platform that can be used for power analysis and optimization without the need for any hardware.

System Modeling Concept

Accurately modeling the inter-coupling nature between different engineering disciplines and simultaneously simulating the constructs on a single platform is crucial for understanding the system. There are high-end tools such as Saber and System-Vision from Cadence and Mentor graphics, respectively, which allow multi-engineering discipline systems to be modeled and simulated on a single platform, but they require specialized knowledge in all relevant engineering disciplines. The cost of licensing these tools is also high and disproportionate to most project costs. Computing power requirement is also intense.

SPICE is a computer modeling language, which allows equivalent circuit models of different engineering disciplines to be modeled in the electronic domain. LTSpice is a free SPICE-based simulation tool developed by power product company Linear-Technology. It allows models to be inputted with their relevant parameters and simulated simultaneously on a single simulation platform. This method requires minimal knowledge of other secondary engineering disciplines, and simulation time and computing power are affordable, due to the simplistic algorithms and relaxed parameters.

The LRUT system consists of multi-engineering-discipline constructs such as electrochemical (battery model), electromechanical (transducers - load) and analog-digital mixed-signal electrical components. SPICE-language -based equivalent circuit models have been developed for simulating foreign domain constructs in SPICE-based simulation tools [5], [6]. However, there is a trade off in the accuracy of equivalent circuit model simulation results compared with the real system. Topology-specific models have been used in modeling work for power and functionally sensitive constructs. Presented here is the modeling and related practical work carried out on the load, transmit-circuit and the high-voltage power supply in the associated system that has significant influence on power performance.

The LRUT System Model

The simplified version of the LRUT system model using LTSpice is shown in Figure 2. It includes several constructs such as transmit circuit, transmit/receive transducer array (PZT_Array), test specimen (Pipeline) and receive circuit (Preamp). The system model represents a pulse-echo mode of operation. The pipeline was modeled using a lossy transmission line model. It has an integrated feature (e.g. defect-weld – acoustic impedance mismatch) and the pipe end is terminated at an acoustic impedance equivalent to air. Each transducer in the transducer array is damped with a stainless steel backing block; more details are given in the load characterization section. This LRUT system model not only allows for analysis of the power performance of the system, but also allows port dynamics to be analyzed for signal strength.

Ultrasonic piezoelectric transducers (lead-zirconate-titanate - PZT) are often used as transmit and receive sensors in LRUT techniques. The capacitive nature of these PZTs requires a high-voltage stress (excitation voltage) to force them into oscillation and to achieve a high signal-to-noise ratio. This high excitation voltage signal needs to be short (broadband) in order to achieve better resolution [7-9].

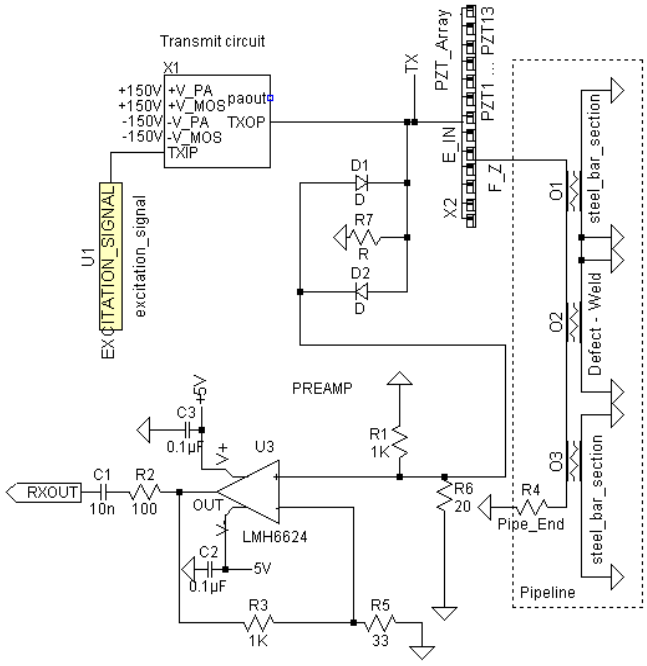


Figure 2. Simplified LRUT System Model Constructed in LTSpice

The LRUT application uses a $240V_{pk\text{-}pk}$ electrical signal in the frequency range of 20kHz to 100kHz. The technique requires a number of excitations (and data collection) at a repetitive rate (rep-rate) of 0.1s for data manipulation. The main circuits involved are the transmit-circuit (TX), receive -circuit (RX), high-voltage power supply (CCPS) and the digital logic control circuits (DSP). The TX circuit produces a high-voltage, high-current electrical signal that excites the transducer array (load) that in turn produces the sound waves. The receive-circuit allows reception and signal processing of the echo signals from features. CCPS is a fast capacitor-charging power supply that produces +/-150V on demand, which provides voltage to the TX. The DSP handles system control, signal processing, storage and communication.

Load characterization

Load charecterisation is required for specifying the PRU's port dynamics and power performance. Load for the PRU is an array of PZT transducers of the type EBL#2 [10] that are pre-engineererd with damping blocks and faceplates. A number of equivalent circuit models for PZT transducers are discussed in the literatue [5]. This work employed a single-dimensional-thickness mode Krimholtz, Leedom and Matthaei (KLM) model, as the KLM model allows additional layers such as face plates and matching layers to be easily added on to the model. Faceplates and matching layers can be modelled using the lossy transmission-line model. The derivation of parameters used in the KLM model for the PZT transducer requires three basic parameters that can only be obtained using practical measurements or by using equations [11]. They are free capacitance (C^T), resonant frequency (f_p) and anti-resonant frequency (f_a) of the transducer. A Solartron SI1260 impedance analyzer was used for the practical impedance analysis. The measured free capacitance was approximately 1100pF at an excitation frequency of 1kHz. The resonant frequency (f_p) and anti-resonant frequency (f_a) were measured as 1.7MHz and 2.4MHz, respectively. Another study claimed that for transducers having a thickness very much smaller than the other dimensions, the vibrations in directions other than thickness are insignificant for modeling purposes [12]. Hence, this single-dimensional model is adequate for the modeling process considered here.

Practical input-impedance analysis results were compared with the simulation results across the frequency range of interest. Figure 3 compares the simulation and practical results obtained for a single PZT. Impedance and phase graphs are set to show 20% and 2% error bars, respectively. A good agreement within 20% was obtained between the simulation and practical results.

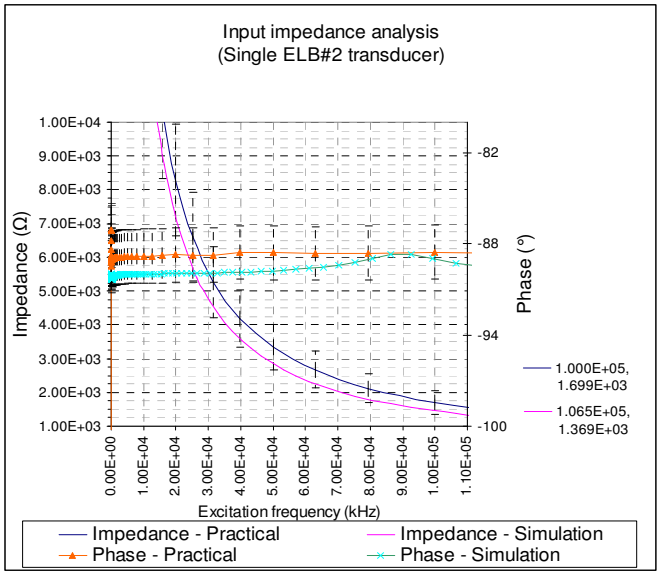


Figure 3. Input Impedance Analysis of a Single Domain PZT

In LRUT applications, PZT elements are mounted to stainless steel backing blocks for damping and mounting purposes. This transducer fabrication also includes a faceplate for acoustic impedance matching and durability. The PZT transducer with backing block and faceplate is called an LRUT transducer. Each output port in the PRU system is specified to drive an array of LRUT transducers. The array size can be as big as 13 LRUT transducers connected in parallel.

The total input impedance analysis for an array of 13 LRUT transducers was also carried out practically and through computer simulations. Faceplates were modeled using a transmission-line model. The stainless steel dampers were modeled with resisters, whose values were calculated using the acoustic impedance formula, $R = \rho A u_p$, where ρ, A and u_p were the density of stainless steel, cross sectional area and phase velocity, respectively [13]. The practical and simulated results are shown in Figure 4. Discrepancies within 30% were observed between practical and simulation results. As the operating region of the LRUT application was well below the series resonance frequency of the PZTs, the load held capacitive properties as expected [11]. This can be seen in Figure 3, where the phase angles are around negative 90 degrees (-90°).

A maximum of 40% variation in input capacitance was observed when practical tests were carried out on two batches of 77 transducers (within and between the batches). Hence, the 30% discrepancy observed in Figure 4 was acceptable. It was concluded from the modeling work that the minimum value of load impedance was 115Ω±30% (80.5Ω), which was confirmed through practical results.

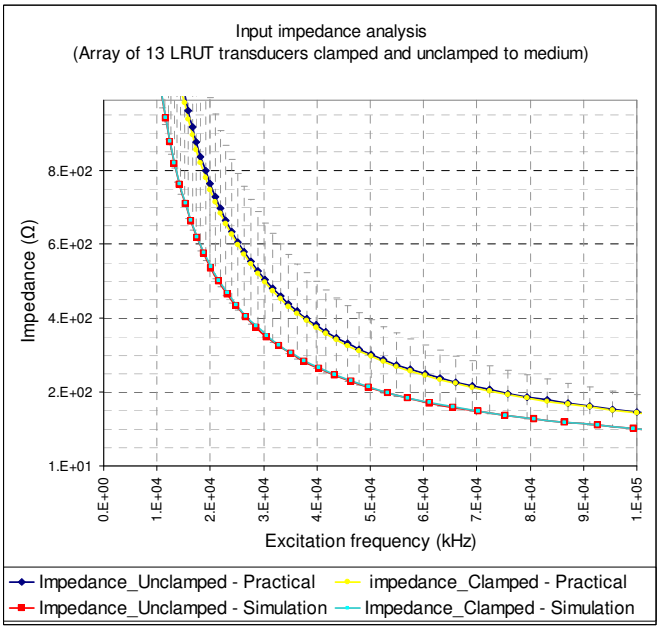


Figure 4. Input Impedance Analysis of an Array of 13 PZT

Transmit Circuit

A transmit circuit capable of exciting a capacitive load with an equivalent impedance of 80.5Ω (70% of the 115Ω for the minimum stated above) was modeled, simulated and prototyped. Using the TX circuit specifications shown in Table 1, design constraints such as slew rate, peak load current and the instantaneous power consumption of the circuit were calculated as 75.3V/µs, 1.5A and 178W, respectively [14], [15].

Table 1. Transmit Circuit Specification

Parameter	Symbol	Range
Tx Supply Voltage	+Vs and -Vs	±150Vdc
Excitation signal frequency	F_{exc}	20kHz-100kHz
Load Value	Z_{total}	115Ω±30%
Excitation Voltage	V_{exc}	±120Vpk
Input signal	Vin	1Vpk-pk
Inverting fixed gain	G	150(43dB)
Number of sine waves per transmit envelop	N_{cycle}	Max 20 Min 10

A complementary MOSFET arrangement, as shown in Figure 5, can provide the required high slew rate, high-voltage excitation waveform and load current demand. The circuit arrangement consists of a single, high-voltage power amplifier (PA) and high-voltage N-type and P-type metal-oxide semiconductor field-effect transistors (MOSFET). The complementary MOSFET arrangement acts as a current source and provides the required load current that the high-voltage PA cannot provide alone. Steele and Eddlemon [16] provide a detailed operation of the circuit. Resistor R_{gs} was chosen to guarantee the maximum required V_{gs} (obtained from MOSFETs datasheets) for the MOSFETs with the PA output current limit [15]. The added auxiliary circuits provide circuit and load protection. The combination of U2-Q1-Rcl+ and U3-Q2-Rcl- provides the current limit protection in the event that the load current exceeds the maximum load current of 1.5A. R6 is a high-value resistor, which provides additional protection for the PA (limits the PA output current), should the MOSFETs open. D1 and D2 are zener diodes that limit the V_{gs} to the maximum specified V_{gs}.

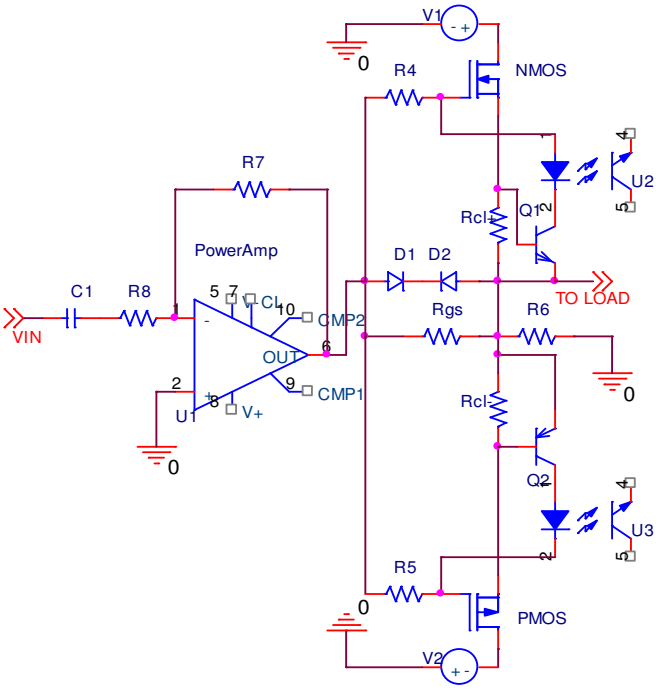


Figure 5. Transmit Circuit

A number of simulations were carried out on this circuit to evaluate stability, SR and, most importantly, power consumption and dissipation at the component level. Simulation results showed that approximately 10% of the supply voltage drops across the MOSFETs and the PA during excitation, generating heat, thus requiring heat sinks. Figure 6 shows the power performance of the circuit for the heaviest load (80.5Ω). It can be seen that the peak instantaneous power consumption of the transmit circuit is about 200W at the excitation frequency of 100kHz for the aforementioned

load characteristics. High current and voltage spikes at the rising edge of the waveforms are due to the short-circuit behavior of the capacitive load at the initial stage (before charging). The quiescent current of the PA was about 10mA [15]; for the applied potential of $300V_{pk\text{-}pk}$, the quiescent power consumption of the circuit was about 3W. This is also noticeable in Figure 6. The simulation results for power consumption were higher than the calculated value (177W) as the calculation steps did not account for power consumption and losses in the auxiliary components. The results presented in Figure 6 also show that the load current was 1.5A at a frequency of 100kHz.

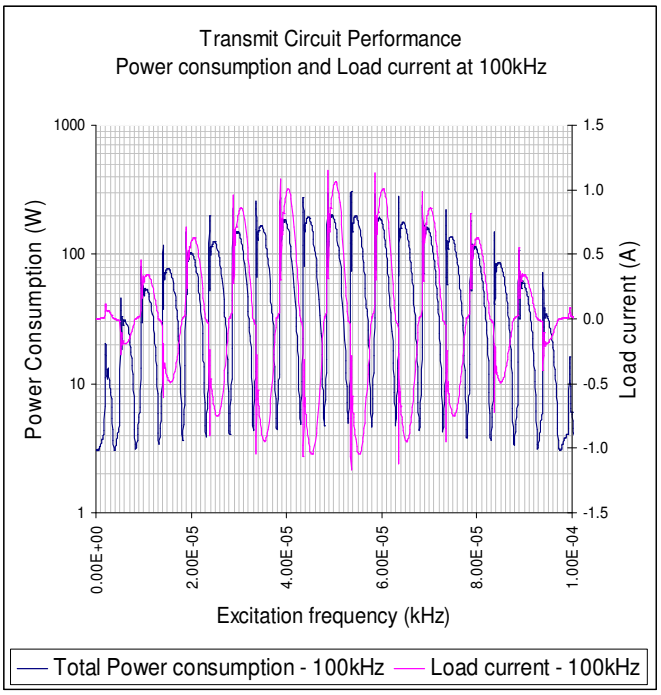


Figure 6. Transmit Circuit Performance

The transmit envelope maximum duration was calculated as 1ms using Equation (1), the values for which are given in Table 1.

$$TX_{duration} = N_{cycle} / F_{exc\_min} \quad (1)$$

The implication of Equation (1) is that the energy utilized by the transmit circuit per transmit cycle is 200mJ (200W x 1ms). A typical PRU can support 24 of these loads, hence 24x 200mJ is consumed immediately. A steady power supply is, therefore, required to provide 200mJ, while maintaining the power rails at +/-150V (low ripple) for each transmit cycle.

High Voltage Power Supply

Normally, power to this type of pulsed load is provided using bulk capacitor banks. In this scenario, energy is stored in capacitors and discharged to the load upon demand. Topping off the capacitor bank is required in order to maintain the voltage between repetitive load pulses. A typical PRU requires separate capacitor banks for +150V and -150V rails with 1,600µF, totaling 3,200µF.

Push-pull converter topology and flyback topology are commonly used in rapid capacitor-charging processes. Push -pull topology is generally used in applications where power requirements are above 200W. A flyback topology was selected for this application due to its simplicity, size, low cost and its widespread use in power applications requiring power levels below 200W [17]. A single-stage flyback CCPS was modeled for meeting the specification of charging a 2x1600 µF capacitor bank within 2s of initial demand and keeping the +/-150 V supply regulation within 4% of target.

The operation of flyback power supply is explained in a report by Basso [17]. The only difference between that work and this study is the split power-supply design. This was achieved using a center-tap transformer. Figure 7 depicts the simplified schematic diagram of the dual–rail, single-stage flyback converter circuit. In flyback topology, transformer T1 is used for maximum energy storage purposes. Hence, it is built with air gaps in the core to trap the energy in them.

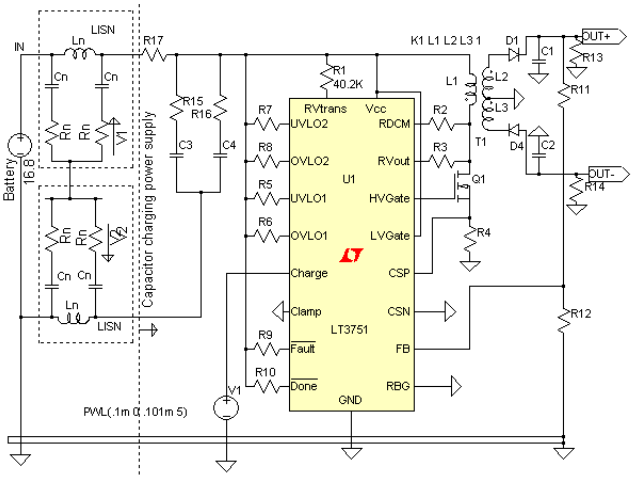


Figure 7. CCPS based on Flyback Topology

When switch Q1 is turned on, the current in the primary inductor (L_1) ramps up with the flux storing magnetic energy. A voltage is induced across the secondary winding (L2, L3) of a polarity such that D1 and D4 are reverse-biased. Hence, no current flows in the secondary circuit. When Q1 is turned off, the primary current drops to zero and the voltage across secondary windings L1 and L2 reverses, allowing D1 and D4 to conduct current that, in turn, charges capacitor banks C1 and C2.

Flyback topology can be used in either discontinuous current mode (DCM) or continuous current mode (CCM). In general, low–voltage, high-current applications use CCM and high-voltage applications use DCM. As this is a high-voltage circuit (transmit current is only drawn from the capacitor banks), DCM is appropriate, but this method is less efficient and tends to heat up the magnetics and the switching device [17]. Boundary mode (BM) operation is also appropriate for this application. In BM, the switching current drops to zero as soon as the switch opens, which make the CCPS very efficient.

There are a number of components which need to be selected before carrying out topology-specific modeling and simulation of this CCPS. They were calculated and listed in Table 2 [18]. A correct selection is important in order to achieve the specified performance.

Table 2. Component Rating for CCPS

Component	Constrains
Transformer inductor	Primary Inductance 11μH; turns ratio 1:10
Switching device N-Type MOSFET	V_{BR} > 32V; $I_{d\text{-}average}$ > 0.94A
Output diodes – Rectifiers	V_{RRM} = 320V or better; $I_{F(AV)}$ > 200mA
Peak Primary current (I_{PRI})	4A

The power-source protection circuits in the PRU are specified to handle the maximum of 10A. The PRU has other circuits that require considerable amounts of current while the capacitor bank is charging. Hence, the primary peak current (I_{PRI}) that the CCPS would use during initial charge up was limited to 4A using R4 in Figure 7.

Simulation results that demonstrate the model fitness of the CCPS are presented in this section. Figure 8 shows the V_{ds} (drain-source voltage) across the switching device, Q1, where I_d is the drain current switching through Q1 when Q1 is turned on. It is clear that the drain of the MOSFET is experiencing an approximate 60V spike due to stray inductance and high di/dt at turn off. The selected MOSFET was rated at 150V, 5A (pulsed current). The drain current, I_d , passing through Q1 when switched on was limited to 4A in the calculation, though the simulation showed that I_d could reach a peak value of 4.2A.

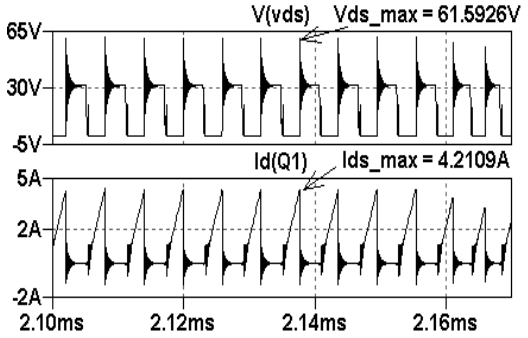


Figure 8. Simulated Switching Device Performance

Figure 9 shows the current passing through the output diode (D1). Average and peak diode current values of 64mA and 240mA were observed. This was expected as calculations indicated that the peak diode current is approximately I_{PK}/2N, or 200mA. The diodes selected for this application can handle 1A peak current. A fast Fourier transform (FFT) of the Q1 switching signal was also probed during simulation to find the maximum switching frequency.

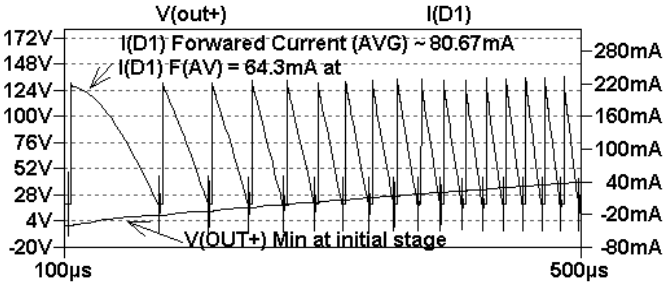


Figure 9. Simulated Rectifier Diode Current

The results are presented in Figure 10 and show Q1 as being switched on at a maximum switching frequency of 150kHz. This value was also used to specify the flyback transformer. For a 20W design, the switching frequency can be between 100kHz and 200kHz [18].

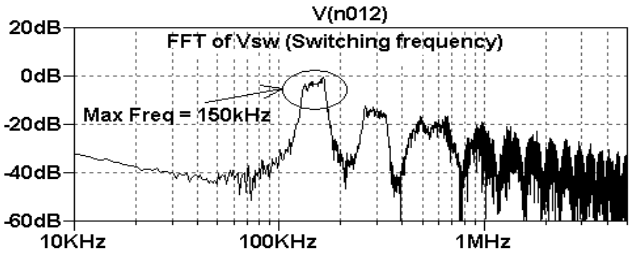


Figure 10. FFT of Switching Signal: Q1 Gate Drive Signal

The amount of electromagnetic interference (EMI) this switch-mode power-supply circuit introduces to the PRU can affect the ultrasonic performance of the system. It is necessary, then, that any EMI within the frequency spectrum from 20kHz to 100kHz would need to be kept below a noise value of -72dB. Differential-mode (DM) and common -mode (CM) noise were evaluated in the simulation using a line impedance stabilization network (LISN) at the input and output (not shown). Differential-mode noise is generally generated due to high di/dt (fast switching) effects on stray inductance. Common-mode noise is generally due to dv/dt effects. This can propagate through the PCB tracks. Localized filtering and careful PCB layout are necessary for noise suppression. Figure 11 depicts the FFT of DM (green trace, top) and CM (blue trace, bottom) noise level at the LISNs inserted at the input stage. Similar work was carried out on the output stage as well (before the capacitor bank). Appropriate filtering was included to prevent EMI spreading to sensitive TX and RX circuits.

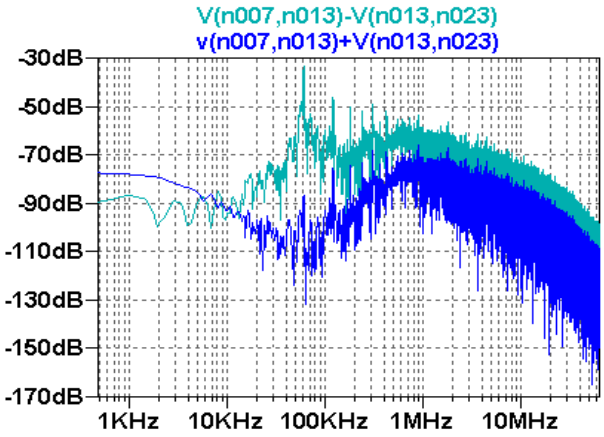


Figure 11. FFT of Differential and Common Mode Noise

Figure 12 highlights the BM-mode operation of the efficient flyback topology. As can be seen, the drain current of Q1 (I_d) drops to zero with the Q1 switching signal (V (n012)), i.e., in BM the primary inductor current (I_{PRI}) drops to zero as soon as the switch (Q1) opens, allowing efficient transfer of energy. The efficiency of the CCPS circuit was analyzed through simulation for varying input voltages (battery voltage) from 13V to 16.8V in 1V steps. Consistent efficiency of 87% was observed for both light and heavy loads.

Practical Validation

Practical results for evaluating the load model were discussed in the load-characterization section. A prototype of the modeled CCPS was produced and practical tests were carried out. Some experimental results are presented to show the validity of the model developed here.

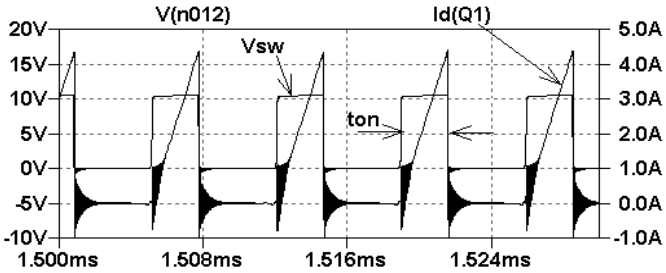


Figure 12. Simulated Boundary Mode Operation of CCPS

Figure 13 depicts the practical measurement taken on the prototyped CCPS for Q1, drain current I_d (CH1/R4), output voltages +150V (CH3), -150V (CH4) and the switching signal (CH2). I_d was measured as high as 18A initially and then dropped to 4A during charging as set by R4 in Figure 7. The initial 18A surge was of short duration and can be tolerated by Q1. In simulation, initial I_d values were seen as high as 26A (not shown in this paper). Maximum switching frequency was measured as high as 179kHz, as shown in Figure 14.

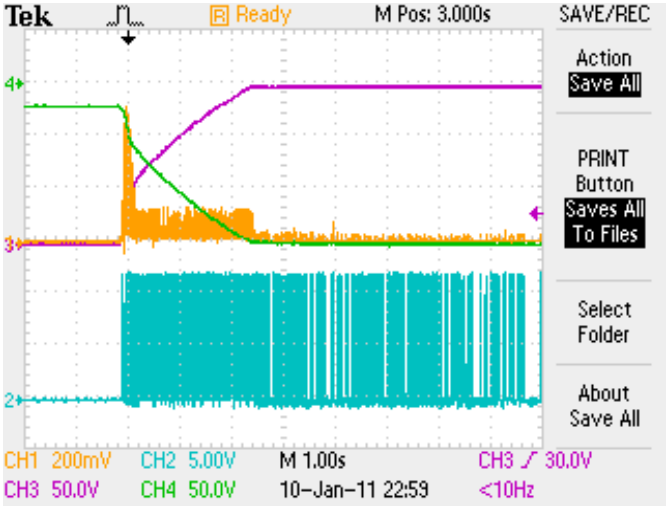


Figure 13. Performance of CCPS - Measured

The voltage stress that the switching device, Q1, experienced (V_{ds}) during switching off and the mode of operation of the CCPS were probed and are portrayed in Figure 15. As can be seen (CH4), the drain voltage, V_{ds}, peaks at 60V when Q1 switches off. The simulation results depicted in Figure 8 also predicted similar values for V_{ds}. The choice of a 150V breakdown voltage for the MOSFET (Q1) was because it could handle this voltage stress. The need for a snubber circuit and the unnecessary power dissipation in the snubber resistor can be avoided by carefully selecting the

device and keeping the PCB tracks short. BM operation is maintained as the drain current of Q1 (CH1) drops down to zero when Q1 is turned off (switching signal - CH2). Other CCPS-related practical results validating the model were obtained, but are not included here.

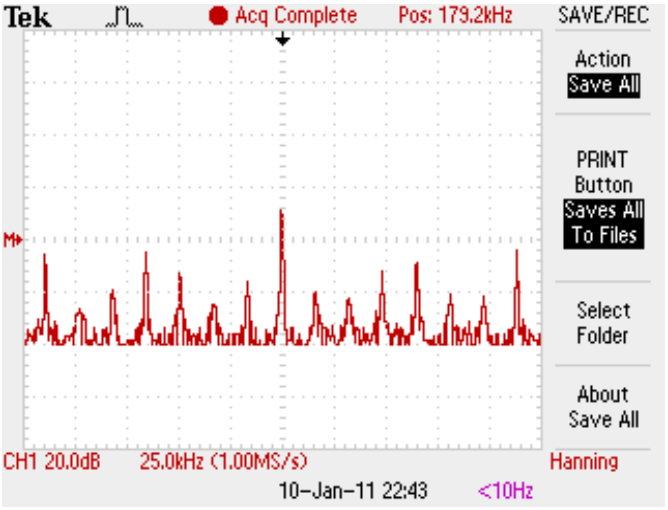


Figure 14. Maximum Switching Frequency of Q1

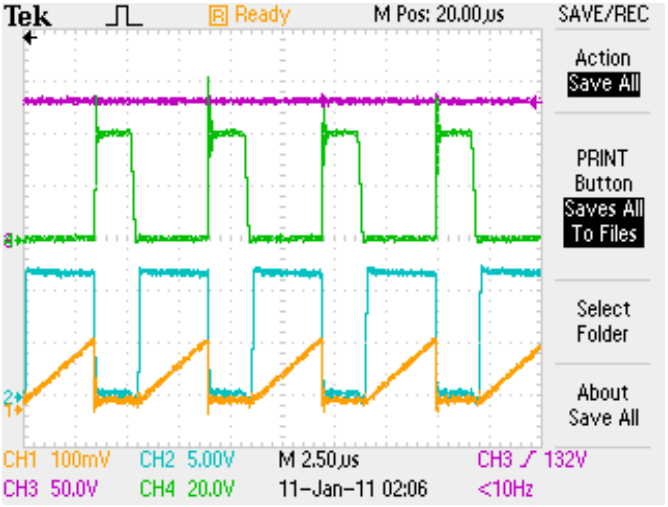


Figure 15. Performance of Q1 and BM Operation

The transmit circuit excitation signal of voltage +/-140V at 50kHz is shown in Figure 16, revealing no cross distortion or slew rate limitation. A voltage drop of 20V across the PA and the MOSFETs of the transmit circuit (Figure 5) was noticed when a +/-150V excitation pulse was demanded at the transmit circuit output, thereby causing voltage clipping at +/-140V.

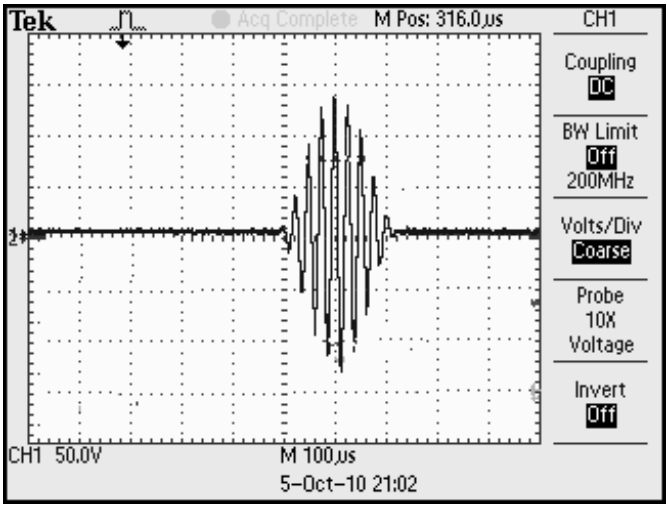


Figure 16. Transmit Circuit Output - Measured

Power Budgeting

A typical power-consumption profile for the modelled LRUT system for one complete inspection cycle (test) was predicted using simulation. A SPICE-based equivalent circuit model of a Li-Ion battery pack was used as a power source for the simulation [6]. The simulation data was used to obtain the adequate power-source capacity value. For both simulation and live measurements, Figure 17 shows how the power source (battery) terminal voltage drops according to the number of tests. The cut-off terminal voltage of a series-connected 4-cell 3.3V (nominal volts) Li-Ion battery pack is about 12V, meaning that the maximum number of complete tests that can be carried out before the likelihood of system failure due to lack of power is 12.

Hardware Realization

A commercial product was manufactured based on the modeling and the prototype and launched at the American Society for Nondestructive Testing (ASNT) spring conference in March 2011 (in Houston, Texas). The production version of the PRU and the model of the internal layout—showing the integrated rechargeable Li-Ion battery—are portrayed in Figures 18(a) and 18(b), respectively. Practical tests carried out on the manufactured PRU revealed that a fully charged power source allowed 20 complete inspection cycles. This satisfied the predicted performance of 12 inspection cycles stated above (an extra 50% battery capacity was added in the production version as a conservative measure to allow an extra 8 inspection cycles for a total of 20).

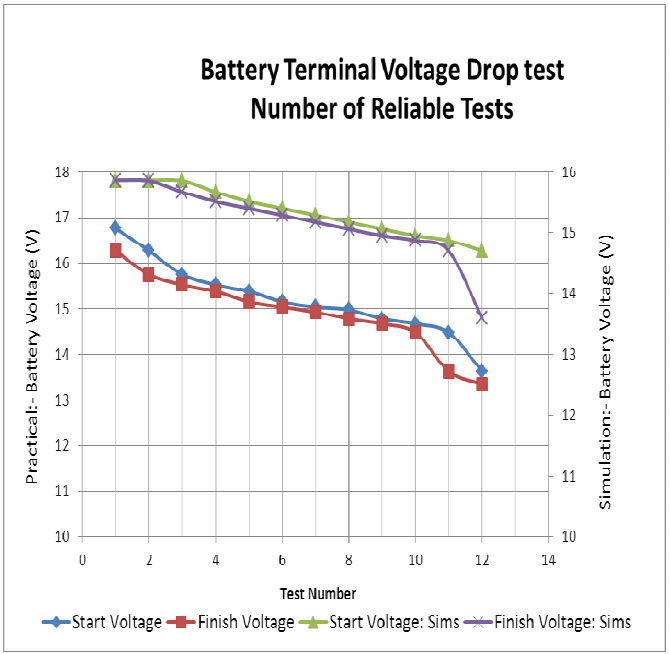


Figure 17. Battery Terminal Voltage versus Number of Tests

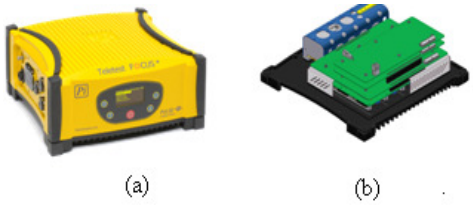


Figure 18. Picture of PRU Product Launched

Discussion

A systematic approach to system modeling allowed rapid prototyping of this industrial application product. It also allowed the developers to avoid over engineering and respinning of the design concept. Close mapping of the simulation results to the hardware results provided a reliable, virtual test platform that industry can use for further enhancement of the product family and foreseeing uncertainties. This research and engineering application work differs from previous work as the system model developed here allows all relevant constructs, regardless of their engineering disciplines, to be integrated and simulated in a single electrical domain simulation platform providing a unique contribution to knowledge.

References

[1] Mudge, P. & Catton, P. (2006). Monitoring of engineering assets using ultrasonic guided waves. *Proceedings of the 9th European Conference on Non-Destructive Testing*. Berlin, Germany.

[2] Catton, P. (2009). *Long range ultrasonic guided waves for the quantitative inspection of pipelines*. Doctor of Engineering thesis, School of Engineering, Brunel University, London.

[3] Akyildiz, F., Weilian, S., Sankarasubramaniam, Y. & Cayirci, E. (2002). A survey on sensor networks. *IEEE communications magazine*, 40(8), 102-114.

[4] Wei, A. (2007). The expected energy efficiency consumption of wireless distributed sensor network based on node random failures. *Proceedings of the communications and networking conference, CHINACOM'07*, China.

[5] Sherrit, S., Leary, S. P., Dolgin, B. P. & Bar-Cohen, Y. (1999). Comparison of the mason and KLM equivalent circuits for piezoelectric resonators in the thickness mode. Proceedings of the IEEE Ultrasonics Symposium, 2, (pp. 921–926).

[6] Gold, S. (1997). A pspice macromodel for Lithium-Ion batteries. *12th battery conference on application and advances*, Long beach, CA.

[7] Schroder, A., Hoof, C. & Henning, B. (2009). Ultrasonic transducer interface-circuit for simultaneous transmitting and receiving. *Proceedings of the ICEMI conference*.

[8] Silva, J. J., Wanzeller, M. G., Farias, P. A. & Neto, J. S. (2008). Development of circuits for excitation and reception in ultrasonic transducers for generation of guided waves in hollow cylinders for fouling detection. *IEEE transactions on Instrumentation and Measurement*, 57(6,), 1149-1153.

[9] Paul, H. (1989). *The Art of Electronics*. (2nd ed.). Cambridge: Cambridge University Press.

[10] Product datasheet for Piezoelectric Precision. (2010). Retrieved November 29, 2010 from http:// www.eblproducts.com/leadzirc.htm

[11] Piezoelectric properties of ceramic materials and components - part 2, BS EN 50324-2, 2002.

[12] Arnau, A. (Ed.). (2004). *Piezoelectric Transducers and Applications*. Germany: Springer.

[13] Marshall, W. (1994). Controlled-source analogous circuits and spice models for piezoelectric transducers. *IEEE transactions on ultrasonics, ferroelectrics and frequency control*. 41(1), 60-66.

[14] Driving Capacitive Loads APEX-AN25. Retrieved November 29, 2010, from http://www.cirrus.com/en/pubs/appNote/Apex_AN25U_F.pdf.

[15] High voltage power operational amplifiers. Retrieved November 29, 2010, from http://www.cirrus.com/en/pubs/proDatasheet/PA90U_K.pd.

[16] Steele, J. & Eddlemon, D. (1993, June). Use high-voltage op amps to drive power MOSFETs. *Electronic design - Analog Applications*.

[17] Basso, C. P. (2008). *Switch-Mode Power Supplies*. New York: McGraw-Hill.
[18] High voltage capacitor charger controller with regulation. Retrieved November 29, 2010, from http://cds.linear.com/docs/Datasheet/3751fb.pdf.

Biographies

T. PARTHIPAN is an Engineering Doctorate (EngD) student in Environmental Technology in the School of Engineering Design at Brunel University in London. He received his BEng (Hons.) degree in Electronic and Electrical Engineering from the University of Surrey and MSc degree in wireless communications from Oxford Brookes University. His interests involve analog and power electronics. He may be reached at billy.parthipan@twi.co.uk

P. MUDGE has been involved in R&D for NDT since 1976 and is currently Technology Fellow in the NDT Technology Group at TWI Ltd, Cambridge UK. He is a Fellow and past-President of the British Institute of NDT and is a Fellow of the Institute of Materials, Minerals, and Mining. He also holds the post of Technical Director of Plant Integrity Ltd and is a Professor Associate in the School of Engineering and Design at Brunel University. He may be reached at peter.mudge@twi.co.uk

R. NILAVALAN is a senior lecturer and course director of wireless communications at Brunel University in London. He received his PhD from Bristol University in 2001. His main research interests include antennas and propagation, microwave circuit designs, numerical electromagnetic modeling and electronic circuit designs. He has published over 80 papers and articles in international conferences and journals in his research area. He may be reached at Rajagopal.Nilavalan@brunel.ac.uk

W. BALACHANDRAN is Professor of Electronics Systems and Director of the Centre for Electronic Systems Research in the School of Engineering & Design at Brunel University, UK. He is a Fellow of IEEE, IET, InstPhy, InstMC and RSA. Professor Balachandran received his BSc degree from the University of Ceylon, Sri Lanka in 1970, and the MSc and PhD degrees from the University of Bradford, UK in 1975 and 1979 respectively. He may be reached at Wamadeva.Balachandran@brunel.ac.uk

INSTRUCTIONS FOR AUTHORS: MANUSCRIPT REQUIREMENTS

The INTERNATIONAL JOURNAL OF MODERN ENGINEERING is an online/print publication, designed for Engineering, Engineering Technology, and Industrial Technology professionals. All submissions to this journal, submission of manuscripts, peer-reviews of submitted documents, requested editing changes, notification of acceptance or rejection, and final publication of accepted manuscripts will be handled electronically. The only exception is the submission of separate high-quality image files that are too large to send electronically.

All manuscript submissions must be prepared in Microsoft Word (.doc or .docx) and contain all figures, images and/or pictures embedded where you want them and appropriately captioned. Also, for all accepted manuscripts, each figure, image or picture that was imported into your Word document must be saved individually as a **300dpi or higher JPEG (.jpg)** file and submitted separately; your manuscript and figure numbers must be used in the title of the file (e.g., **M12-S-18 Figure 4**); that means one additional file for each image imported into your manuscript. These 300dpi images do NOT need to be embedded in your manuscript. For tables or graphs created directly in Word, you do not need to submit them as separate files.

Included below is a summary of the formatting instructions. You should, however, review the sample Word document of our website for details on how to correctly format your manuscript. The editorial staff reserves the right to edit and reformat any submitted document in order to meet publication standards of the journal.

The references included in the References section of your manuscript must follow APA-formatting guidelines. In order to help you, the sample Word document also includes numerous examples of how to format a variety of scenarios. Keep in mind that an incorrectly formatted manuscript will be returned to you, a delay that may cause it (if accepted) to be moved to a subsequent issue of the journal.

1. **Word Document Page Setup**: Two columns with ¼" spacing between columns; Top of page = ¾"; Bottom of page = 1" (from the top of the footer to bottom of page); Left margin = ¾"; Right margin = ¾".

2. **Paper Title**: Centered at the top of the first page with a 22-point Times New Roman (Bold), Small-Caps font.

3. **Page Breaks**: Do not use page breaks.

4. **Body Fonts**: Use 10-point Times New Roman (TNR) for body text throughout (1/8" paragraph indention); 9-point TNR for author names/affiliations under the paper title; 16-point TNR for major section titles; 14-point TNR for minor section titles; 9-point TNR BOLD for caption titles for tables and figures; other font sizes as noted in the sample document.

5. **In-text Referencing**: List and number each reference when referring to them in the body of your document (e.g., [1]). The first entry must be [1] followed by [2], [3], etc., continuing in numerical order to the final entry in your References section. Again, see the sample Word document for specifics. Do not use the End-Page Reference utility in Microsoft Word. You must manually place references in the body of the text.

6. **Tables and Figures**: Center all tables and figures. Captions for tables must be above the table, while captions for figures are below; all captions are left-justified.

7. **Page Limit**: Manuscripts should not be more than 15 pages (single-spaced, 2-column format).

8. **Page Numbering**: Do not use page numbers.

9. **Publication Charges:** Manuscripts accepted for publication are subject to mandatory publication charges.

10. **Copyright Agreement:** Review the **copyright transfer agreement**.

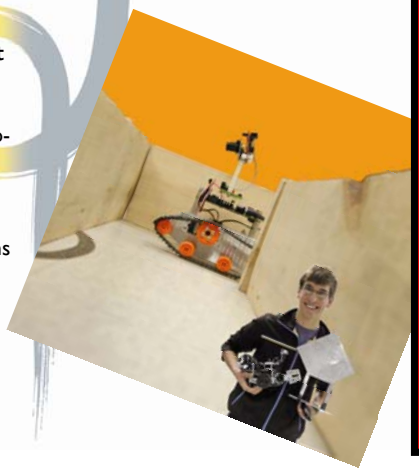
11. **Submissions:** All manuscripts and associated files must be submitted electronically; the only exception would be if you wish to submit a CD with the individual files for your high-quality images, as noted above.

    MANUSCRIPTS should be submitted to Dr. Philip D. Weinsier, manuscript editor, at philipw@bgsu.edu along with a copy to editor@ijme.us.

    FILES containing your high-quality images should ONLY be submitted to philipw@bgsu.edu.

College of Engineering, Technology, and Architecture

University *of* Hartford



For more information please visit us at www.hartford.edu/ceta

For more information on undergraduate programs please contact Kelly Cofiell at cetainfo@hartford.edu.

For more information on Graduate programs please contact Laurie Grandstrand at grandstran@hartford.edu.

Toll Free: 1-800-766-4024
Fax: 1-800-768-5073

DEGREES OFFERED :

ENGINEERING UNDERGRADUATE

Acoustical Engineering and Music (B.S.E)
Biomedical Engineering (B.S.E)
Civil Engineering (B.S.C.E)
-Environmental Engineering Concentration
Environment Engineering (B.S.E)
Computer Engineering (B.S.Comp.E.)
Electrical Engineering (B.S.E.E.)
Mechanical Engineering (B.S.M.E.)
- Concentrations in Acoustics and
- Concentrations in Manufacturing

TECHNOLOGY UNDERGRADUATE

Architectural Engineering Technology (B.S.)
Audio Engineering Technology (B.S.)
Computer Engineering Technology (B.S.)
Electronic Engineering Technology (B.S.)
-Concentrations in Networking/ Communications and Mechatronics
Mechanical Engineering Technology (B.S.)

GRADUATE

Master of Architecture (M.Arch)
Master of Engineering (M.Eng)
- Civil Engineering
- Electrical Engineering
- Environmental Engineering
- Mechanical Engineering
  - Manufacturing Engineering
  - Turbomachinery

3+2 Program (Bachelor of Science and Master of Engineering Degrees)
E^2M Program (Master of Engineering and Master of Business Administration)

College of Engineering, Technology, and Architecture
UNIVERSITY OF HARTFORD

IJME NOW SEEKS SPONSORS

IJME IS THE OFFICAL AND FLAGSHIP JOURNAL OF THE INTERNATIONAL ASSOCATION OF JOURNALS AND CONFERENCE (IAJC)

www.iajc.org

The International Journal of Modern Engineering (IJME) is a highly-selective, peer-reviewed journal covering topics that appeal to a broad readership of various branches of engineering and related technologies. IJME is steered by the IAJC distinguished board of directors and is supported by an international review board consisting of prominent individuals representing many well-known universities, colleges, and corporations in the United States and abroad.

IJME Contact Information

General questions or inquiries about sponsorship of the journal should be directed to:

Mark Rajai, Ph.D.
Editor-in-Chief
Office: (818) 677-2167
Email: editor@ijme.us
Department of Manufacturing Systems Engineering & Management
California State University-Northridge
18111 Nordhoff St.
Northridge, CA 91330