Following the Paper Trail:
Reliable Processing of Voting Records for Trustworthy Elections

December 2006

Proposal funded by the National Science Foundation under award numbers NSF-0716368, NSF-0716393, NSF-0716647, NSF-0716543. Any opinions, findings, and conclusions or recommendations expressed on this website are the investigators' and do not necessarily reflect those of the National Science Foundation.

Investigators

Professor Daniel Lopresti
Department of Computer Science and Engineering
Lehigh University
Bethlehem, PA

Professor Ziad Munson
Department of Sociology and Anthropology
Lehigh University
Bethlehem, PA

Professor Christopher Borick
Department of Political Science
Muhlenberg College
Allentown, PA

Professor George Nagy
Department of Electrical, Computer, and Systems Engineering
Rensselaer Polytechnic Institute
Troy, NY

Professor Elisa Barney Smith
Department of Electrical and Computer Engineering
Boise State University
Boise, ID

Collaborative Research: CT-T: Following the Paper Trail: Reliable Processing of Voting Records for Trustworthy Elections

Section B. Project Summary

What was once seen as an expensive but straightforward transition to electronic voting is now viewed with increasing suspicion as a situation rife with potential risks and vulnerabilities. Computer security experts have stated unequivocally and with near unanimity that paper can play a fundamental role in guaranteeing safe and secure elections, either in the form of hand- or machine-marked ballots, or as a Voter Verified Paper Audit Trail (VVPAT). However, the processing of such records during the initial counting of votes or in the conduct of recounts has raised its own set of problems which span technical and social boundaries.

We propose to address the issues that currently make paper records more of a nuisance than an integral component in trustworthy voting systems. Specifically, we plan to characterize the statistical distribution of mark sense errors as a function of ballot quality, create a system for unbiased visual auditing based on ballot images and study its effectiveness, examine the possibility that a concept known as homogenous class display (HCD) can facilitate manual recounts, consider whether new ID technologies such as RFID and digital watermarking have a role to play in securing paper ballots, evaluate error behavior when VVPAT is used with Direct Recording Electronic (DRE) systems, analyze the user interface issues that arise in interactions between touch screen selection, paper records, and ballot / audit trail design, and develop testing procedures for the paper handling components of voting systems in accordance with operational constraints, including the modest training received by most poll workers. Our work on voting technologies will be supported by – and supports – a series of extensive user studies, surveys, and focus groups that we plan to run to measure voter understanding and acceptance.

Intellectual Merit

The novel focus of this proposal is the integration of previous research in a variety of technical areas – human-machine interfaces, pattern recognition (specifically optical character and optical mark recognition), and image processing – with public perceptions of technologies used in governance. The in-depth examination of the effects of ballot characteristics on human and machine interpretation that we propose is long overdue. Our planned study of the statistical dependences between errors that can lead to biased election results, the several methods that we propose to eliminate such bias, and our ideas for applying human cognitive abilities in an election context precisely when and where it is indispensable, are all firsts of their kind.

Broader Impact

It is difficult to conceive of anything with broader impact than the development of error-free and trustworthy voting technology. Because elections are governed by laws and procedural rules promulgated at multiple levels of government, and of the preparations required by the sheer size of many elections and the cost of change in terms of equipment purchase, voter education, and the training of election officials, voting technologies tend to evolve slowly. Right now, however, we are at a cross-roads, with several radically different voting technologies competing for acceptance. It is therefore timely to examine the human and cybernetic dimensions of the major contenders, with a focus on the role that paper records have to play. Our research should also benefit difficult current transitions in processing financial transactions and educational tests.

Collaborative Research: CT-T: Following the Paper Trail:
Reliable Processing of Voting Records for Trustworthy Elections
PROJECT DESCRIPTION

1. Introduction

Florida's infamous “butterfly ballot” and resulting recount efforts during the 2000 presidential election, leading directly to the Help America Vote Act (HAVA), have initiated the most dramatic change in the way America votes in the history of our country. Not long after the subsequent introduction of electronic voting equipment, computer security experts and concerned citizens began raising serious questions concerning the reliability and trustworthiness of such systems when collecting, storing, and tabulating votes [KSRW04, RISC04, H05, B06, H06, FHF06]. Direct Record Electronic Voting (DRE) systems, once seen as a straightforward albeit expensive solution, is increasingly being regarded as an unacceptable compromise.

In its recent draft report on Voluntary Voting Systems Guidelines for 2007 [BKPW06], the Security and Transparency Subcommittee for the Technical Guidelines Development Committee of the National Institute of Standards and Technology (NIST) wrote:

> “One conclusion drawn by NIST is that the lack of an independent audit capability in DRE voting systems is one of the main reasons behind continued questions about voting system security and diminished public confidence in elections. NIST does not know how to write testable requirements to make DREs secure, and NIST's recommendation to the STS is that the DRE in practical terms cannot be made secure. Consequently, NIST and the STS recommend that VVSG 2007 should require voting systems to be of the SI [software independent] “class,” whose readily available (albeit not always optimal) examples include op scan and DRE-VVPAT [Direct Record Electronic with Voter Verifiable Paper Audit Trail].”

The NIST report continues on to observe that the use of paper to provide independent auditing capabilities in elections is entirely practical, but that there are undeniably open technical issues that can and should be addressed:

> “The widespread adoption of voting systems incorporating paper did not seem to cause any widespread problems in the November 2006 elections. But, the use of paper in elections places more stress on (1) the capabilities of voting system technology, (2) of voters to verify their accuracy, and (3) of election workers to securely handle the ballots and accurately count them. Clearly, the needs of voters and election officials need to be addressed with improved and new technology. The STS believes that current paper-based approaches can be improved to be significantly more usable to voters and election officials ...”

As it stands today, paper is treated as a simple “add-on” when it should be an integral component in trustworthy voting technology. A good deal of attention is already being paid to other parts of election systems; in this proposal, we elucidate the substantial benefits that accrue when paper artifacts are accorded the attention they deserve.

Introducing (or re-introducing) paper into the election process is not above criticism, some of which is certainly valid and must be addressed. Eliminating or mitigating current drawbacks associated with hardcopy form the basis for our proposed research. We also note that there are a myriad of other, non-technical concerns that have been voiced, ranging from a feeling that paper is “old fashioned,” to worries about increased “litter,” and the added cost of the required consumables (paper, ink/toner for printers, etc.) Another major thrust of our project will be to conduct comprehensive studies of citizen attitudes toward voting technologies, the various ways in which paper might be incorporated into the election process, and establishing trust for the tabulating and auditing capabilities of various schemes now in use or newly-developed as a result of our work. To our knowledge, ours will be the first studies of these kinds.

A recent article in the New York Times [UD06] hits on a number of the key points:

> "New federal guidelines, along with legislation given a strong chance to pass in Congress next year, will probably combine to make the paperless voting machines obsolete, the officials say. States and counties that bought the machines will have to modify them to hook up printers, at federal expense, while others are planning to scrap the machines and buy new ones ...
>
> Advocates for the disabled say they will resist his bill, because the touch-screen machines are the easiest for blind people to use ...
>
> Paul S. DeGregorio, the chairman of the federal Election Assistance Commission, which was created by Congress in 2002 to set voting standards, also cautioned against rushing to make changes, especially since some counties also ran into problems with printers in this year's elections ...
>
> Beverly Kaufman, the county clerk [in Harris County, TX], said she and other election officials elsewhere disliked the paper requirement ... "Every time you introduce something perishable like paper, you inject some uncertainty into the system," Ms. Kaufman said."

The evidence that major hurdles lie ahead is growing rapidly. In another recent article, the Atlanta Journal-Constitution [C06] reports:

> "An examination of the much discussed e-voting paper-trail audit system used in several Georgia districts revealed that the process is far from a silver bullet. The paper-trail solution gained popularity in the wake of many warnings from computer scientists that e-voting machines were vulnerable to tampering. Cobb County, Ga. head of elections Sharon Dunn recently testified in front of state officials, who are considering making paper-trials mandatory in all of the state's voting districts, regarding the effort needed to manually count the 976 printouts generated in the district: Twenty-eight people took part in the five-day task of counting the votes from 42 races, and teams often had to restart their counts as numbers did not match up. Dunn testified, "It looks easy until you have to do it." ... MIT political science professor specializing in elections Charles Stewart said, "Audits ask humans to do something that computers are generally better at doing."

It is our premise that this popular attitude – that paper-based voting records are inherently more problematic and less reliable – is fundamentally incorrect. Nonetheless, there are undeniably problems that must be addressed. For example, paper jams can wreak havoc when DRE (Direct Recording Electronic) systems are equipped with printers to produce a paper audit trail. Cuyahoga County, Ohio, found that a manual count did not match the computer-tallied results because 10 percent of ballots were either smeared, torn, crumpled, or blank [M06]. In Chester County, PA, a close election that would determine the majority party in the PA House of Representatives was disputed when one party insisted that a recount be conducted by running the optical scan ballots through a different brand of scanner hardware, noting that the tallies can vary depending on the system in use [WHPTV06].

The certification examinations for electronic voting systems in Pennsylvania are a matter of public record [PA06] and are videotaped. During the recent certification of one optical scanning system [PAV06], the final vote tallies for nearly all of the candidates of one political party were wrong, while the votes for the candidates of the other party were all right. There is no reason to suspect the scanner was functioning incorrectly – the system provided no indication that anything was amiss. In the end, a decision was reached that the problem was due to some unspecified hardware/software "bug" and the system was certified for use. The presence of such failure modes in electronic voting systems, completely undetectable during normal operation, is one of the urgent issues we plan to address.

Moreover, it is our thesis that trustworthy elections are a function of the hardware and software used to collected, tabulate, and audit votes, as well as the experience and attitudes of the citizens who use the system (or decide not to). E-voting protocols shown to be mathematically "sound" through a rigorous theoretical analysis might still fail to win public acceptance if the voting process is regarded as incomprehensible, the threats are not well understand by average citizens, or the procedures for auditing elections and conducting recounts do not inspire confidence.

Our interest in public perceptions of voting technologies pre-dates our preparations for this proposal. In a professional telephone survey we conducted last fall [LM06], 523 Pennsylvania voters expressed moderate levels of confidence in the new e-voting systems, but they also expressed overwhelmingly support for machines that show voters paper verification of the votes they cast, with 81% stating they believe such verification is important. In comparing the public's trust in electronic voting with their trust in other widely-used technologies, voting machines came out somewhere in the middle. Overall, our respondents trust electronic voting machines much less than ATM machines, but more than financial transactions on the Internet or even airport security systems. These intriguing results just begin to scratch the surface of an issue at the very core of our democracy, and we plan to extend and enhance our measurement of public attitudes and trust over the course of our project.

We intend to examine the options presented by three alternative uses of hardcopy:

(1) voter-verified paper audit trail (*VVPAT*) output by direct-recording electronic (*DRE*) systems,

(2) human-marked ballots counted by hand or via image processing techniques, and

(3) machine-marked ballots which likewise can be counted either way.

While many computer security experts argue strongly in favor of options (2) and (3), option (1) may be popular with municipalities that have already invested heavily in current DRE e-voting systems. Some believe it may be possible to upgrade such systems with the addition of a printer in a way that mitigates certain of the shortcomings in the current products.

In option (2), we will concentrate on more efficient and effective methods of counting and auditing paper ballots, both at the voting location and at election headquarters. However, the most promising and interesting possibilities may lie with option (3), which spans both the reading and interpreting of ballots from current screen-based ballot-marking devices (essentially computer-controlled printers), to the development of a new *direct-action* electro-optical marking device, to be described in Section 3, that supports immediate verification, local and remote vote counting, and both automated and human auditing.

In the next section, we detail our plans to study these issues and develop solutions to increase the reliability and security of elections as well as public confidence in the results.

2. Proposed Experimental Work on the Reliable Processing of Voting Records

We have cited a range of problems that are inherent in the hardware and software design of current optical scanning systems, where the investigators on this proposal have decades of expertise [GN81, GN83, GN92, LNT96, ZLSN97, ES98, SQ04, SLZN98, RNG99, GN00, GN04, LN05, SA06]. Known image processing methods can be used to minimize undetected failure. We can devise tests for detectable failures, develop real-time diagnostics, warn the poll workers and election judges about such failures, and provide guidelines for ballot layouts that ensure random or systematic errors do not disadvantage one candidate over another. Our past work shows such errors are not uniformly distributed across the page [ELS94, LTZZ95, LZ97, ZL94].

The main focus of our work – and the set of tasks we are proposing – is on the handling of paper voting records, whether ballots or hardcopy audit logs. Paper brings with it certain fundamental benefits. The goals, then, are finding ways to process it with high accuracy, to integrate it into election machinery at reasonable cost, to eliminate or reduce the chances for paper-based fraud, and to enable minimally trained individuals to both use it (as voters or administrators) and trust it. Our proposed project is not an attempt to design a complete, end-to-end e-voting system, but our results will be applicable to a broad range of e-voting systems that produce some form of paper artifact. Neither will we compete with the experienced research teams who are already developing new cryptographic methods for providing election security, or with those studying formal methods for software engineering or verification. Below we outline the experimental tasks we propose to perform on various types of hardcopy records.

Task 1. Statistical distribution of mark sense errors as a function of ballot quality

Regardless of whether a ballot was marked by hand or by machine, one critical output variable of a vote counting/auditing system based on mark sense technology is the accuracy of mark recognition per *mark*, per *ballot*, and per *machine*. These numbers, which can be readily estimated experimentally, are already available from past studies of both electoral ballots and multiple-choice tests, from manufacturers, and from elections having undergone recounts. Such data is, however, virtually useless without a thorough understanding of the underlying joint probability distributions that reveal the temporal and spatial statistical dependences between errors [GN95, ZLSN97, SLZN98, GN04]. In order to obtain this information, we will first conduct pilot experiments with a small number of ballots specimen printed with different printers and in different colors [NS04]. Once we are confident that we can produce ballot specimens that adequately reflect the variability and range of both hand-marked and machine marked ballots, we propose to print, using a high quality laser printer, a statistically significant number of ballots with a geometrical distribution of mark locations that reflects current ballot design.

Our experimental design calls for repeated scanning of ballots in 3 layouts (*tight*, *medium*, *loose*), and 5 qualities (determined by the *extent* to which the mark fills the designated bullet, and its *contrast* and *texture*). We will print 1,000 ballot specimens (with a total of about 100,000 marks) of each of the 15 designs, and run them 10 times through modern mark sense readers used in elections. Each ballot will also have a mark-sense specimen identity code, printed at the highest quality with sufficient redundancy to ensure negligible error rate. At 1 second per ballot, testing a single reader will require 15x10x1000 =150,000 seconds = ~40 hours. We believe that during the first year, we can test 3-4 optical mark scanners.

We intend to freeze the statistical methodology for analyzing results *before* further experiments are performed. Statistical analysis of the results (multi-variable regression and analysis of variance of errors vs. scan order and mark positions) will allow us to draw conclusions towards formulating standards for *paper format* and *composition* (stock, dimensional stability, thickness, grammage, stiffness, bursting/tearing strength, opacity, porosity, smoothness), *ink/toner characteristics*, *printer mechanism* (ribbon, dot matrix, laser, inkjet), *mark design* (symbology, layout, fiducial marks), *human-readable ballot content* (layout, drop-out colors, font, type size), and *scanner technology* (paper transport, illumination color and source, detector, and optics).

Our major focus here is determination of correlations between the digital image and the mark-sense errors. Current scanner systems lack a robust reject-error trade-off. According to our OCR experience, there is no technical reason for tolerating multiple mis-scans instead of rejecting the entire ballot for manual processing.

For each specimen, we will record (1) the identity of the form, (2) the results of the mark-sense output, and (3) a complete digital image at 300 dpi (to be used in further experiments described below). Fortunately the image data is sparse. At a 100:1 lossless compression ratio (well within the capabilities of JBIG-2, and even of the earlier CCITT-G4, JBIG, and DjVu), each megapixel (~8x10x300x300) ballot image is about 10 KB, or a total of 1.5 GB for each scanner under test. We will make these ballot images available for study to other interested researchers as well.

The mark sense technology used today was developed decades ago and is largely based on binary image processing. With current scanner technology, there is little cost, speed, or storage benefit associated with using binary scans. We will therefore also investigate the intrinsic trade-off between *depth* (number of gray levels) and s*patial sampling rate* (pixels per unit area) for ballot scanners, as well as improvements for *gray-scan binarization*, based on existing algorithms. As a byproduct of the above experiments, we will gather data on the *throughput* (peak and average, per machine), *failure rate*, and *failure mode* of each machine, keeping track of fixed and operating *cost* and physical scanner characteristics (*size, weight*, and *vulnerability to tampering*).

Task 2. Unbiased visual auditing based on ballot images

Election fraud could be carried out by corrupt vote counters who add spurious marks to ballots, thereby invalidating them (because of conflicting marks) in subsequent recounts. If instead votes are counted on *displays* of digital ballot images, this scam is rendered essentially unworkable. There is still, however, the possibility of voluntary or involuntary bias in counting. We intend to conduct two types of experiments that may lead to simple ways of avoiding such bias. Both will make use of the ballot images collected in Task 1.

The first set of experiments will consist of comparing human vote counts *using context* versus human counts *without using context* (oddly called *blind* recording). Context is present when enough of the ballot is visible to the human counter to identify the party, candidate, or proposition for which a vote was cast. Context is absent when only the mark area is displayed. We intend to conduct this experiment with a digital display, but clearly the same context-removal procedure could be implemented mechanically on paper ballots using a predesigned template. Since our specimen ballots do not carry candidate labels, these will be generated and displayed synthetically. We will instruct our experimental judges to count the votes accurately, but for the contextual alternative we will construct labels that induce some subconscious bias.

Semi-automating the processing of election recounts could address one of the most pressing (and well-publicized) issues we face in moving to paper support for electronic voting. While the Remote Encoding System (human encoding of postal address images) has been an economic success for the United States Postal Service, we must take into account the additional legal and social constraints imposed in the case of elections.

Task 3. Homogenous class display

The second experiment designed to verify automated mark sense counts is inspired by long-established methods of OCR verification, Homogenous Class Display (*HCD*). Here the isolated images where a mark was registered will be grouped for display. The images where a mark was *not* registered will be similarly grouped. The positive marks will be displayed in groups of 400 (20x20) simultaneously, and the human operator will be asked to enter any discrepancies. The negative (i.e., blank) instances will be displayed in similar groupings, and again the operator will notate any perceived discrepancies. Spurious marks will be introduced to assess human error rate. (In some OCR service bureaus, the operators are notified of the presence of artificial errors in order to keep them alert.) HCD is, in principle, a very fast and convenient method of verifying automated counts, but it can never be accepted without experimental evidence of its accuracy.

Following the practice of the postal RES, we intend to conduct variations on HCD that include displaying only bullets that were not classified as *marked* or *blank* with very high confidence.

Experience with high-accuracy financial OCR and MICR (magnetic ink character recognition) indicates that reject/error ratios of the order of 100:1 or even 1,000:1 are necessary to achieve acceptable levels of undetected errors at the election precinct level. Since errors within a ballot are correlated, a more sophisticated reject strategy can be based on the confidence level of multiple marks on the same ballot.

Using machine-printed (albeit imperfect) ballots avoids tedious ground-truthing. But we can still compare human accuracy on the paper ballots themselves with human accuracy on their images as a function of *ballot quality*, *scan resolution*, *inspection speed, display size,* and *display quality*. The resulting information should be extremely valuable for both election planning and commercial development of improved voting machinery.

Auditing elections through electronic images of the paper trail offers certain benefits over auditing the paper trail itself. The main advantage is that image-based auditing can be done by physically separated election inspectors in parallel. Furthermore, it is infinitely repeatable because the electronic image is preserved intact. There is an added cost for a limited amount of write-once storage for each scanner, though. Whether such auditing would prove acceptable to voters is a key question that we plan to examine on the social sciences side of our project.

These experiments are intended to demonstrate that the proposed procedures can guarantee the absence of *systematic* bias, and to determine how many duplicate inspections are required to reduce the expected number of *random* errors to any prespecified level that is deemed acceptable.

Task 4. Analysis of synthetic ballot images

We plan to purposefully scan our ballots at much higher spatial sampling rate than is necessary for mark-sensing in order to be able to construct lower-resolution images by sub-sampling. Lowering image resolution can result in less expensive and faster devices and also further decrease storage cost, but this cannot be done safely without quantitative understanding of the resolution/error trade-off. Since some research groups advocate injecting artificial noise for stress-testing of election machinery, we seek an opportunity to test whether, contrary to our earlier conclusions, artificial noise provides a viable test procedure [LNT96].

Typical document image quality measures are based on the distribution of the shape and size of the foreground components. We will develop a measure of *ballot quality* based on our earlier research on statistically detectable local distortions [ES98, ES01a,b,c, SQ02, SQ04, ES05].

Task 5. Unique identification of ballots

Several classical voting scams are based on ballot switching which can take place on the way to the voting station, when the ballots are collected, or during a recount. Although in recent years physical security measures and custody-chain policies appear to have largely eliminated en-masse ballot substitution, we will revisit our earlier proposal of embedding radio frequency identification (RFID) in ballots [DLGN05], and compare it to watermarking and 2-D bar-coding. Some current RFID devices, such as the Hitachi mu-chip, are already thin enough to be embedded in medium-weight paper, and their cost is approaching that of a single sheet of paper.

We do not envisage any experimental work associated with this task, because RIFID is usually marketed as a complete system, but if paper with embedded RF transponders does become commercially available during the grant period, we will assess its benefits for ballot identification. We envision the simplest RFID system, where a single unique 12-digit number is associated with each ballot prepared for an election precinct and recorded at the time the ballot is printed, and the list of numbers separately transmitted to election officials. If the blank ballots are handled by the voters rather than automatically fed to a remote-control mark printer, then it may be desirable to introduce a further read station in the local scanner.

The objective of this task will be finding ways of guaranteeing voter anonymity without giving up the security benefits of the identifiably of the ballots themselves.

Task 6. Error characteristics of DRE with VVPAT

We note that markings created by a human should be minimal in order to save time. No voter (or test-taking student) would want to darken three bubbles instead of one in order to record a single bit of information! This same constraint does not, however, apply to either machine-generated marks or to machine-generated alphanumeric symbols. It takes only a very short code to represent a single bit of data extremely reliably: even a simple five-bit majority code lowers the error rate by three orders of magnitude. We will therefore devise appropriate means of encoding votes intended for OMR (optical mark recognition) or OCR.

In DRE with VVPAT, the voter has an opportunity to verify that the touch-screen device recorded the vote correctly. This is done by inspecting a printed output generated by a separate device attached to the touch screen, much like the verification of a charge-card transaction through a printed receipt. The voter cannot, however, obtain a copy of the printed record of the vote, because this may allow payment for a prespecified vote ("vote selling").

A major drawback of the roll-paper record provided by some e-voting vendors is that it preserves not only each vote, but also the voting sequence. This could be compared to the observed arrival of voters and imperil voter anonymity. Most such systems therefore print only enough information (at the beginning and the end of a voting session) to support a limited form of tallying from the paper record, recording the full audit trail on a write-once medium. Many questions regarding the adequacy of such audits remain to be resolved, however. Assuming that cut-paper printers with full vote records (a la the "Mercuri Method" [M05]) are substituted for roll-paper printers, then we may consider such a device similar to a machine ballot marking system. The major difference is that here the vote is still recorded electronically, and the paper record may never be scanned. (Whether the legal ballot is the electronic data or the paper record is a question we leave to legislatures to decide.)

Since such systems may be adopted by some election districts, we will examine the error rate of the OCR system necessary to process the human-readable record [GDAS05]. On the basis of our earlier experience, we believe that the required accuracy can be reached with (1) an appropriate OCR font, (2) adequate vertical and horizontal spacing [ES06], (3) self-inking paper or carbon-film ribbon, (4) coded input, and (5) minimal printer maintenance. Experiments to verify this conjecture will be conducted with synthetically generated alphanumeric voting records over a range of qualities similar to those of the mark-sense records described in Task 1.

Task 7. Analysis of touch screen interfaces

In terms of user accessibility, touch screen interfaces offer certain benefits over more traditional methods of marking ballots. Indeed, this point was one of the major driving forces of the HAVA legislation. While some user interface issues have already been examined in specific contexts (e.g., the NY State Board of Elections study of the voting time users required in a mock election [NP06], the reports on residual voter error rates by the Caltech/MIT Voting Technology Project [CM01], a recent survey by George Tech researchers of voters with disabilities [BRMW05]), our interests lie in the interactions between the touch-screen portion of the system, comprising an integrated graphical display and a mechanism for registering user inputs, and the paper records printed by either a ballot-marking device or the VVPAT when the voter casts his/her votes. We plan to measure the standard critical variables, including voting time, voter confidence, required training, voter and device error rates, and system cost, durability, and robustness.

We will also examine a variety of user interface issues as they pertain to e-voting systems that produce hardcopy records, including the positioning and rendering of candidate names (choice of font, font size, colors, layout of the ballot / audit trail), feedback to confirm votes, mechanisms for correcting mistakes as well as for confirming or cancelling ballots, and implications for access by the disabled (Section 5). Although we are not planning a general study of touch screen voting, we will characterize complete systems where paper records are an integral component.

Task 8. Development of testing procedures for different types of voting systems

Our review of the procedures for certifying e-voting systems leads us to believe that test protocols that already fall short of providing the necessary assurances will break down completely for the systems we propose to process paper voting records at ultra-low error rates. New estimation techniques will be required, as will operational procedures for differentiating voter error from machine error, determination of optimal "reject" rates for mark sense systems, sampling methods for comparing the results of electronic and paper tallies, statistical methods for uncovering suspicious runs of votes that may indicate a latent hardware or software malfunction, and the insertion of predetermined test votes into the voting stream to catch errors as early as possible. Our experience in bank check and postal address reading, two high volume, high accuracy applications, suggests several quality control measures appropriate for election systems.

The impact of "wear and tear" on voting systems is different than in financial, postal, and standardized test-grading applications because of the preponderance of minimally trained poll workers and local officials. In contrast to research and commercial settings, voting machines are often kept locked in dusty, humid storerooms most of the year. We plan to develop tests to detect and counter the resulting malfunctions.

Beyond inhospitable storage environments, software bugs, and malicious attempts to compromise the system, there is no doubt that human error is responsible for a great many problems. While voter error rates have been studied fairly extensively (e.g., in the reports on voter error rates by the Caltech/MIT Voting Technology Project), the impact of errors made by poll workers and local election officials have received less attention. Again, our focus is specifically tuned to systems that incorporate paper records. How can poorly-trained and perhaps technology-hostile workers inadvertently mess things up? Can better design mitigate such problems? We plan to assemble the data needed to answer such questions.

3. New Voting Technologies

In addition to our studies around the use of paper in existing systems, we plan to construct a prototype of a new type of voting machine that combines the advantages of the *direct action* [EEB67] of a touch-screen device with the auditability of a paper ballot. It exploits the properties of pressure-sensitive or thermal-transfer markers, such as some ribbon label makers. We tentatively call our device the *Lincoln*.

With the Lincoln, the votes are marked on a transparent touch sensitive screen, with a paper ballot directly behind it (untouchable by the voter) on which the visible marks appear. A computer records the position of each touch (without knowing what candidate it is for!). The touch also marks the ballot. The ballot immediately enters a scanner integrated with the ballot transport. The scanner records the position of the marks. It also records the entire ballot image on a write-once device (with lossless compression). Optionally, both the scanned image and the mark reader results can be displayed immediately to the user on an ancillary screen.

The ballot is dropped into a bag and shipped to the county seat. Optionally the ballots are uniquely identified by RFID or watermark or barcode. The counts from both the microprocessor that interrogates the screen, and the optical mark reader, are recorded on write-once media and also shipped to the county seat. Different jurisdictions can decide on different control procedures, such as comparisons of the:

1. Polling station screen-based counts
2. Polling station mark reader counts
3. Another pass through the ballot *image* by a different mark reader
4. Rescan and mark sense of the hardcopy ballot
5. Human inspection of the ballot image (cf Section 2)
6. Human inspection of the original paper ballot

We have devised various schemes for correcting errors due to pressing of the wrong location on the screen (errors are immediately visible by the darkening of ink on the ballot). These methods are all based on adding, rather than erasing, marks.

The Lincoln would offer several advantages over current systems. Paper jams are less likely because the integrated design has a fixed paper path, and the paper is introduced into the scanner mechanically. The voter does not carry the ballot to a scanner, therefore the process is faster and less floor space is needed. Security is higher, since none of the Lincoln microprocessors know the set-up of the ballot for the particular election: the counts are recorded only by ballot position. (A feature somewhat similar to Chaum's punchscan system [P06], although realized in a different way.) Voter anonymity is protected. Errors by election officials are less likely, because initialization is simpler. Only the paper ballots must be printed correctly.

We believe that the Lincoln satisfies many of the opposing arguments for *and* against paper trail. Since everything is within sight and the interaction is "natural," voter confidence will be enhanced. Such an integrated system could also be price-competitive in a few years.

We have the electronic, mechanical and human factors expertise required to construct a prototype Lincoln. While some questions remain, the major steps facing us appear to be:

(a) Selection of the transparent touch screen materials
(b) Selection of touch-sensitive ballot stock
(c) Design of the paper transport behind touch screen
(d) Circuit design of the mark location electronics (E)
(e) Overall microprocessor control design
(f) Mechanical support and integration of a standard roll-feed optical scanner
(g) Specification of the Graphic Voter Interface (GVI)
(h) Specification of ballot-layout constraints
(i) Evaluation of voter reaction and trust in Lincoln

4. Voter Understanding, Trust, and Acceptance (Task 9)

Our research on voter attitude will extent the survey we conducted of Pennsylvania voters prior to the 2006 midterm elections, and also address a noteworthy result of a recent voter study on 812 subjects that was sponsored by the New York State Board of Elections [NP06]. The main objective of the latter was to establish the MDR (Maximum Daily Rate) for seven different systems and different user populations. The report includes demographics and some data on the handicapped and on use of "assist" features. While not the focus of this study, one question attempted to measure voter trust in the technology: "I am confident that the voting system accurately recorded my vote." Surprisingly, confidence levels were lower for optical scan systems than for purely electronic (DRE) voting: while voters marked their own paper ballots and watched them feed through the scanner, they were not provided with any confirmation that their votes are being read correctly by the scanning software. As we noted earlier, optically marked ballots are not always read correctly, and without receiving feedback, users are right to question whether their votes have been counted. "Blind" scanning is an issue we plan to address in our research. Another important dimension that the NY study does not address, but that we consider central to our work, is how easy people believe it would be to ignore / change / rig an election. This dimension is especially important when it comes to electronic systems because such a large proportion of the population sees such technology as a "magical black box."

The social science research in many ways bookends the computer science / engineering experiments. Up front, we propose to carry out two national surveys. The first would be a nationally representative survey of approximately 1,000 voters regarding their overall experience with different voting technologies, levels of trust in those technologies, and general attitudes toward the political process. The purpose of the survey is to obtain a baseline average for the voting public as a whole on a series of measures of attitudes and experiences toward voting and voting technologies, and trust in technological innovation and the political process. Although several limited studies have been conducted to date, including our own on Pennsylvania voters, this information does not yet exist at the national level.

The second national survey would be much larger (~4,800 voters) with respondents drawn from subpopulations that vary on a) the type of voting technology used in their county, b) the length of time the county has been using the current technology, and c) the level of overall political competitiveness in the county. These subpopulations are necessary in order to evaluate variations in public acceptance of voting results as well as how and why Americans place their trust in some voting technologies and not others.

The smaller national study will be conducted directly by the investigators on the project with extensive experience in attitudinal research and survey design, using facilities already in place at their home institutions. The second, larger, study will be designed and directed by the investigators, but the actual data will be collected by well-known scholarly survey firms.

We see this survey work as the front end of the project because it will provide rich data on (a) how the public views voting technology, (b) what particular aspects of the technology they are more and less comfortable with, and (c) how the technology can be improved to increase overall levels of public trust in the systems. These findings will then feed into the computers science aspects of the project. The findings will define the boundaries of what the public will accept, how they evaluate and interpret the technical processes behind counting votes, and generate some ideas about what kinds of technology they would embrace or resist.

The second phase of social science work anchors the other end of the project. After we develop several models of what secure systems might look like and how they would operate (like the Lincoln prototype described above), we would conduct a series of several dozen focus groups with adult registered voters in the Lehigh Valley of Pennsylvania. We chose the Lehigh Valley not only for convenience, but also because the region is broadly representative of the US population as a whole on a range of politically relevant variables (such as income and education levels, relative numbers of Democrats and Republicans, etc.) [A04,WHH04]]. The region thus offers an excellent laboratory for studying public reaction to changes in voting technology.

The investigators will introduce different voting technologies and different election outcome scenarios to participants in order to gauge both the usability of the systems and — more importantly — their reaction to the machines and their trust that the systems can accurately record their votes and resist attempts at fraud. Sessions will expose participants to mock-ups of proposed interfaces, log their activity, and allow them to discuss their reactions to the systems in open discussions facilitated by principal investigators in this study along with their graduate assistants. On the computer science side, these experiments will reveal the overall time and consistency of voter interactions with the systems when facing different interfaces, as well as discover whether they spend more time on some aspects of the ballot (e.g. propositions or higher office) than others. We can also determine whether they are equally likely to correct all errors (which could be artificially introduced). On the social science side, these experiments will expose the underlying mechanisms of attitude development toward different voting technologies responsible for the patterns of trust we explore in our two national surveys.

The composition of the focus groups will reflect the voting-age population of the United States as a whole. We anticipate that each session will last approximately two hours, and all focus group activity will be video and audio recorded for later transcription and analysis. Our final report will integrate the results of the analysis of voter attitudes with objective measurements of the constituents of the voting process.

5. Voters With Disabilities (Task 10)

Providing voting systems accessible to citizens with disabilities was one of the driving forces behind the Help America Vote Act. Touch screen systems provide several obvious affordances in this regard, including the ability to increase the font size and contrast ratio for voters suffering from visual impairments, the option of incorporating text-to-speech synthesis (TTS) to "read" ballots to voters who are blind, multilingual display options that are set to the voter's preferred language when he/she first approaches the machine, and sufficient portability / flexibility to allow voters in wheelchairs to make use of the system.

We note, however, that the careless incorporation of paper records in the election process can "break" many of these affordances. A VVPAT cannot be verified by someone who cannot see; current systems often print such records on narrow, low-contrast, cheap devices designed for printing retail sales receipts in small and poorly distinguishable characters [RNG99].

An informative study conducted by researchers at Georgia Tech [BRMW05] produced the following conclusions:

- Voters with disabilities were 7% to 20% less likely to report satisfaction with the equipment than voters without disabilities.
- Of the complaints that voters expressed, 20% concerned the display characteristics of on-screen text, 31% were regarding the placement or design of controls, and 48% were due to the quality of the audio / sound output.
- More voters with disabilities reported problems (60%) than voters without disabilities (39%).
- Finally, 16% of voters said that lack of privacy was a key concern in using e-voting machines.

Hence, while HAVA was intended to address the question of accessibility, it is clear that many issues still remain. We have studied accessibility in conjunction with the behavioral aspects of text editors [NE81], and will specialize our methods to the *readability* (content) and *legibility* (appearance) of the short, low-entropy messages involved in voting. Our primary concern will be how the use of paper records in an election setting, either as ballots or as a VVPAT, impacts voters with disabilities. Paper should not be seen as adverse to accessible elections.

6. Educational Outreach

We have had repeated interactions with the Coalition for Voting Integrity (CVI), the League of Women Voters, and other non-partisan civic groups. Daniel Lopresti, for example, was recently interviewed as an expert on e-voting security for an hour-long radio program hosted by CVI [CVI07]. He also organized an open panel session on electronic voting held at Lehigh in April 2006 which was attended by many citizens and covered by the local news media; Christopher Borick also served as a panelist. We will continue our active participation in forums oriented towards election-technology and provide our associates with materials we find in the professional literature (e.g., ACM *Communications* and IEEE *Computer*) that they might otherwise miss.

We will develop educational materials on voting technology, automatic scanning and counting of votes on paper ballots, mark reading, and the types of errors that machines can make, for the Boise State University *e-Girls* program. This program is intended to counter some of the causes of under-representation of women in engineering [NSF96, NSPE92, NRC91, NRC92, GT90, TGB98]. It is held every summer for girls entering 10th and 11th grades to explore the technical relevance between engineering and their personal interests. Past modules have included *Biomechanics of high-heeled shoes; Materials science of fashion fabrics; Physics of rock climbing* and many others. To encourage participation from all socioeconomic backgrounds, e-Girls is provided free of charge with support from private foundations, the Idaho National Lab, and an NSF 2005 PAESMEM award received by BSU engineering dean Cheryl Schrader. The program draws students from a wide geographic area in southern Idaho which has a rapidly growing Hispanic population [I00]. BSU is tracking which students from the first cohort pursue a technical college career in 2007, and several have already committed to engineering.

Likewise, Lopresti selected *Electronic voting and its attendant problems* as the topic for team projects in the introductory engineering course at Lehigh which all first year engineering students must take. This exposure has increased the awareness of students possessing a range of technical interests, including a number of female and minority students. This activity will continue as well. Two juniors working under his supervision (one male, one female) were selected to present their work on paper audit trails for e-voting at a statewide symposium on undergraduate research.

Support for this proposal will enable Nagy to continue to promote high-level participation of women and minorities in science and engineering. Three of his recent doctoral graduates, and the research assistant on his current NSF grant, are women: they have published 25 journal and conference papers related to their doctoral research. Three of his recent women undergraduate research assistants are pursuing graduate degrees He has served on several Women-in-Engineering committees, and also actively mentors minority undergraduates. He believes that education in science must be international, and therefore intends to continue collaborating with universities in Canada, China, France, Germany, Italy, Japan, India, Sweden and Switzerland, and to review research and education proposals for foreign national funding organizations.

As part of this project, we plan to host two broadly-constructed workshops bringing in representatives from the numerous constituencies that have a stake in these issues. We hope to include local, state, and federal election officials, politicians, civics teachers, voters' rights groups, representatives from the handicapped community, "hackers," and fellow academics. Attendance will be both through targeted invitations and open calls for participation.

7. Broader Impact

It is difficult to conceive of anything with broader impact than the development of reliable and trustworthy voting technology. Because elections are governed by laws and procedural rules promulgated at multiple levels of government, the preparations required by the sheer size of many elections, and the cost of change in terms of equipment purchase, voter education, and the training of election officials, voting technologies tend to evolve slowly. Right now, however, we are at a cross-roads, with several radically different voting technologies competing for acceptance. It is therefore timely to examine the human and cybernetic dimensions of the major contenders, with a focus on the role that paper records have to play.

The survey and focus group components of the research will speak not only to bringing the power of new technology to the practical and critically important arena of the democratic process, they also will provide critical new data for ongoing debates over the role of civil society and public trust in the proper functioning of democratic institutions. Even before Robert Putnam published *Bowling Alone* [RB00], which documents the ways in the erosion of social connections in the United States have diminished public trust, scholars and political observers alike have worried about declining rates of voting and public apathy toward politics. Our proposed research will contribute to ongoing efforts to understand the chief mechanisms responsible for these trends and the ways in which public policy can address them.

8. Related Efforts

The project we are proposing is complementary to existing efforts examining various aspects of electronic voting. Below we list some of the better-known activities, noting that there are also many researchers working individually.

ACCURATE: A Center for Correct, Usable, Reliable, Auditable, and Transparent Elections

ACCURATE is a multi-institution voting research center funded by NSF under the CyberTrust program. In a broad sense, the goals of the center – to research ways in which technology can be used to improve voting, to educate the public about voting-related issues, and to serve as a resource to a variety of constituencies – mirror our own. A perusal of the research interests and publications of ACCURATE participants shows work on user interfaces, privacy, testability of voting systems, and new data structures for more secure auditing. The usability of traditional (hand-marked) paper ballots was examined in a study involving human subjects [EBG06]. Not reflected in the center's activities is the sort of work we plan on characterizing the machine processing of paper voting records, developing methods that result in ultra-low error rates, exploring new ways of using paper to facilitate trustworthy elections, and measuring public trust and acceptance of these various alternatives.

Princeton Center for Information Technology Policy

The Center for Information Technology Policy at Princeton University addresses societal issues that go beyond electronic voting, such as privacy and security, that have a connection to advances in computer technology. Their publication of a detailed analysis of the flaws in a touch screen e-voting system produced by a major vendor just before the 2006 midterm election raised public awareness of these issues [FHF06]. Given their broad mandate and the backgrounds of the faculty who are participants, however, it appears unlikely that the Center for Information Technology Policy will attempt to address the sorts of technical issues we plan for our project.

MIT-CALTECH Voting Technology Project

Created in the wake of the 2000 presidential election, the MIT-CALTECH project performs large-scale longitudinal studies of elections for residual votes, comparing the original results and recounts sorted by technology. While such studies are relevant to the work we plan, this effort is not focused on developing new mechanisms for processing voting records as we are proposing.

The UConn Voting Technology Research Center

The UConn VoTeR Center takes as its mission advising state agencies in the use of voting technologies and developing procedures for safely using such equipment in elections. As with the Princeton Center for Information Technology Policy, they performed a detailed analysis of an existing e-voting system (in this case, an optical ballot scanner) identifying numerous security flaws [KMRS06]. It does not appear to fall within the mandate of the VoTeR Center, however, to perform the kinds of research on developing new methods for reliably processing paper ballots and audit trails that we have presented in this proposal.

9. Results from Prior NSF Support

Daniel Lopresti

Daniel Lopresti's current NSF CyberTrust award, "Collaborative Research: Using Generative Models to Evaluate and Strengthen Biometrically Enhanced Systems" (CNS-0430338, 2004-2007), is a joint research effort with colleagues Fabian Monrose at Johns Hopkins and Michael Reiter at Carnegie Mellon. This project examines behavioral biometrics in the face of new threat models, and proposes new evaluation methodologies that more accurately characterize the security of such systems. Among other results to date, we have demonstrated that some adversaries are able to improve the effectiveness of their handwritten forgery attacks with modest amounts of training, and that computer generative models for handwriting can equal the performance of talented human forgers. This work has produced four noteworthy publications over the past two years [BLM06, L06, BML06, LR05], with a journal submission now under second review. The PIs have given several invited talks on the topic, including a keynote address delivered by Lopresti at the First Korea-Japan Joint Workshop on Pattern Recognition (travel expenses were paid by the workshop organizers). Two master's students at Lehigh have been supported so far, both having moved on to positions in the computer software industry.

George Nagy

George Nagy's current NSF grant, Table Analysis for Generating Ontologies (TANGO), Award ID: 0414854 (2005-2008) is his first since 1991. Results from six prior NSF grants were published in archival journals, including Computing Surveys and IEEE Transactions. During the first year of the current grant, an interactive system that improves with use was developed for converting arbitrary web tables to standard Wang notation. The extraction of content cells and their relation to nested header categories is largely automatic. The logging and statistical analysis system necessary to evaluate the improvement over current table-extraction methods is under construction. This project builds on Nagy's, Embley's, and Lopresti's widely-cited earlier work on table analysis and ontology construction.

Elisa Barney Smith

Elisa Barney Smith's work has been funded by the grants "CAREER: Document Image Degradation Analysis" (CCR-0238285: effective 2003-2008), which included two REU supplements. Other NSF funding includes a sub-contract from an NSF EPSCoR grant (9720634) awarded as a Faculty Startup Grant "Document Scanning Defect Analysis Using Bilevel Image Features" (subcontract from NSF EPSCoR grant 9720634: (effective 2000 - 2001).

This work has developed and refined models of the printing and scanning process. A degradation model that describes the spatial density of toner particles on paper was developed [NS04]. A relationship among the different degradations of a bilevel (text) image was incorporated into a method to improve OCR [BSA06, BSA05]. Methods to estimate the degradation model parameters from text images without the presence of specialized test charts were also developed [Y04].

Three full-time graduate students in the Boise State University's Electrical and Computer Engineering Department have been supported under the CAREER grant. Fifteen undergraduates have been funded by this grant to work on aspects of this project during the summer or the school year. Four of them are now pursuing master's degrees, two of whom have continued with this project. The outreach component of the CAREER grant introduced elementary education majors to engineering through an alternative science methods class. This course has been run successfully three times and two papers have been presented on this work [MBS06, BSM05].

10. Coordination Plan

Barney Smith, Lopresti and Nagy have collaborated on several projects over the last dozen years, exchange visits regularly, and expect to continue meeting several times a year during the duration of the project. Borick, Lopresti, and Munson normally see each other every week and jointly developed the 2006 Lehigh / Muhlenberg survey of Pennsylvania voter attitudes. We will make sure that all the graduate students have a chance to meet all the PIs at least once each year, and have each attend at least one conference during their dissertation research. The proposed timeline and the PIs most involved in each task are shown in the diagram below.

Research Timeline and Task Leaders

```
Survey 1      BM      xxx
Survey 2      BM              xxx  xxx
Workshop 1    MLB              xx  xx      x
Workshop 2    MBL                                      xx  xx
Task 1        NSL   xxxxx  xxxxxx  xxxxxx
Task 2        LN      xxx  xxxxxx  xxxxxx  xxx
Task 3        NL             xxxx  xxxxxx
Task 4        SL           xxxxxx          xxx
Task 5        LN                      xxx  xxxxxx  xxxx
Task 6        LS      xxx  xxxxxx  xxxxxx
Task 7        LS           xxxxxx  xxxxxx  xxxxxx
Task 8        LNS                  xxxxxx  xxxxxx  xxxxxx  xxxxxx
Task 9        BM   xxxxxx  xxxxxx  xxxxxx  xxxxxx  xxxxxx  xxxxxx
Task 10       LS                           xxxxxx  xxxxxx  xxxxxx
Focus Groups  BM                                      xxx  xxx
Lincoln       SNL  xxxxxx  xxxxxx  xxxxxx  xxxxxx  xxxxxx
                   YEAR 1          YEAR 2          YEAR 3
```

LEGEND: **B**: Borick; **L**: Lopresti; **M**: Munson; **N**: Nagy; **S**: Barney Smith;

TASKS

1. Ballot imaging and error distribution
2. Unbiased visual audits
3. Homogenous Class Display
4. Synthetic ballot images
5. Unique ballot identification
6. DRE with VVPAT
7. Touch-screen interface
8. Testing procedures
9. Test benches for social science aspects
10. Handicapped access

As noted earlier, we plan to host two 3-day workshops as part of this project, the first during the summer of the second year (2009) and the second during the summer of the third year (2010). We envision that the earlier workshop will focus on technological issues, while the later will deal more with sociological questions and the outcomes of our focus group and user studies.

11. Expected Outcomes

The proposed study will explore citizen attitudes towards technology in an essential societal context. It may lead to new ways of increasing confidence in information processing technology. On the technical side, this research will engender more efficient *lay* human interaction in OCR and MBR in ultra-low error-rate applications similar to the paper trail in e-voting, where interaction will tremain indispensable for decades. The investigation of unbiased human intervention and of improved error-estimation is also likely to benefit other critical applications, as will effective test protocols. The proposed research will help accelerate the drawn-out transition from binary to grey-scale and color scanning for text and mark images, and experience with Lincoln may transfer to more secure terminals for electronic transactions.

References

[A04] "Welcome to Swingtown," Jim Axelrod, CBS News, April 27, 2004 (available at http://www.cbsnews.com/stories/2004/04/27/eveningnews/main614002.shtml; accessed 1/21/06).

[B06] "The Machinery of Democracy: Protecting Elections in an Electronic World," Report of The Brennan Center Task Force on Voting System Security (Lawrence Norden, chair), June 27, 2006. http://www.brennancenter.org/programs/downloads/Full%20Report.pdf

[BKPW06] "Requiring Software Independence in VVSG 2007: STS Recommendations for the TGDC," William Burr, John Kelsey, Rene Peralta, and John Wack, National Institute of Standards and Technology, November 2006. hittp://vote.nist.gov/DraftWhitePaperOnSIinVVSG2007-20061120.pdf

[BLM06] "Evaluating the Security of Handwriting Biometrics," L.. Ballard, D. Lopresti, and F. Monrose, *Proceedings of the Tenth International Workshop on Frontiers in Handwriting Recognition*, October 2006, La Baule, France.

[BML06] "Biometric Authentication Revisited: Understanding the Impact of Wolves in Sheep's Clothing," L.. Ballard, F. Monrose, and D. Lopresti, *Proceedings of the Fifteenth USENIX Security Symposium*, July-August 2006, Vancouver, BC, Canada.

[BRMW05] "Disabilities, (E) Voting and Access: A Survey of Voter Satisfaction," Paul M. A. Baker, Robert G. B. Roy, Nathan W. Moon, and Michael Willaims. http://www.cacp.gatech.edu/Presentations/CSUN_2006/Paul/CSUN_2006_e-vote.pdf

[BSA05] "Text Degradations and OCR Training," E. H. Barney Smith and T. Andersen, *International Conference on Document Analysis and Recognition*, Seoul, Korea, August 2005, pp. 834-838.

[BSA06] "Partitioning of the degradation space for OCR training," E. H. Barney Smith and T. Andersen, *Proc. SPIE Electronic Imaging, Document Recognition and Retrieval XIII*, Vol. 6067, San Jose, CA, January 2006, paper #606705.

[BSM05] "Engineering for Educators," E. H. Barney Smith and R. Miller, *Proceedings ASEE Annual Conference*, Portland, OR, June 2005, paper #9259.

[C06] "Voter Paper Trail Not an Easy Path," Carlos Campos, *Atlanta Journal-Constitution*, December 22, 2006.

[CM01] "Residual Votes Attributable to Technology: An Assessment of the Reliability of Existing Voting Equipment," The Caltech/MIT Voting Technology Project, March 30, 2001.

[CN82] R.G. Casey and G. Nagy, "Recursive Segmentation and Classification of Composite Character Patterns," *Proceedings of Sixth International Conference on Pattern Recognition*, pp. 535-537, Munich, Germany, October 1982.

[CVI07] "CVI on Radio -- Voice of the Voters," Coalition for Voting Integrity, http://mysite.verizon.net/resq4lzq/cvi/id181.html

[DLGN05] D. Lopresti and G. Nagy, "Chipless RFID for paper documents," *Procs. Document Recognition and Retrieval* XII, SPIE Vol. 5676, pp. 208-215, January 2005.

[EBG06] "Measuring the usability of paper ballots: Efficiency, effectiveness, and satisfaction," S. P Everett, M. D. Byrne, and K. K. Greene, *Proceedings of the Human Factors and Ergonomics Society 50th Annual Meeting*. Santa Monica, CA, 2006.

[EEB67] English, W.K., Engelbart, D.C., and Berman, M.L. "Display selection techniques for text manipulation." *IEEE Transactions on Human Factors in Electronics HFE*, 1967, pp. 5-15.

[ELS94] "Classification and Distribution of Optical Character Recognition Errors," J. Esakov, D. Lopresti, and J. Sandberg, *Proceedings of the IS&T/SPIE International Symposium on Electronic Imaging, February 19*94, San Jose, CA, vol. 2181, pp. 204-216.

[ES01a] E. H. Barney Smith, "Scanner Parameter Estimation Using Bilevel Scans of Star Charts," *Proc. International Conference on Document Analysis and Recognition 2001, Seattle*, WA, 10-13 September 2001, pp. 1164-1168.

[ES01b] E. H. Barney Smith, "Bilevel Image Degradations: Effects and Estimation," *Proc. 2001 Symposium on Document Image Understanding Technology*, Columbia, MD, 23-25 April 2001, pp. 49-55.

[ES01c] E. H. Barney Smith, "Estimating Scanning Characteristics from Corners in Bilevel Images," *Proc. SPIE Document Recognition and Retrieval VIII*, Vol. 4307, San Jose, CA, 21-26 January 2001, pp.176-183.

[ES02] E. H. Barney Smith, "Uniqueness of Bilevel Image Degradations," *Proc. SPIE Document Recognition and Retrieval IX*, Vol. 4670, San Jose, CA, January 2002, pp. 174-180.

[ES05] E. H. Barney Smith and T. Andersen, "Text Degradations and OCR Training," *International Conference on Document Analysis and Recognition*, Seoul, Korea, 29 August - 1 September 2005, pp. 834-838.

[ES06] E. H. Barney Smith, "PSF estimation by gradient descent fit to the ESF," *Proc. SPIE Electronic Imaging, Image Quality and System Performance III*, Vol. 6059, San Jose, CA, January 2006, paper #605914.

[ES98] E. H. Barney Smith, "Characterization of Image Degradation Caused by Scanning," *Pattern Recognition Letters*, Volume 19, Number 13, 1998, pp. 1191-1197.

[FHF06] "Security Analysis of the Diebold AccuVote-TS Voting Machine," Ariel J. Feldman, J. Alex Halderman, and Edward W. Felten, Princeton Center for Information Technology Report, September 13, 2006.

[GDAS05] P. Garrison, D. Davis, T. Andersen, E. H. Barney Smith, "Study of style effects on OCR errors in the MEDLINE database," *Proc. SPIE Electronic Imaging, Document Recognition and Retrieval XII*, Vol. 5676, San Jose, CA, January 2005, pp. 28-36.

[GN00] G. Nagy, "Twenty Years of Document Image Analysis in PAMI," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, #1, 20th Anniversary Issue, pp. 38-62, January 2000.

[GN04] G. Nagy, "Classifiers that improve with use," Proceedings of Conference on Pattern Recognition and Multimedia, pp. 79-86, *IEICE*, Tokyo, February 2004 (invited).

[GN81] G. Nagy, "Criteria for Selecting Automatic Digitizers (Optical Scanners)," Harvard Computer Graphics Week, 27 pages, Cambridge, MA, July 1981 (invited).

[GN82] G. Nagy, "On the Frontiers of OCR," *Proceedings of the IEEE*, vol. 40, #8, pp. 1093-1100, July 1992.

[GN83] G. Nagy, "Optical Scanning Digitizers," *IEEE Computer*, vol. 16, #5, pp. 13-25, May 1983.

[GN92] G. Nagy, "What does a machine need to know to read a document?" *Proceedings of ISRI Symp. on Document Analysis and Information Retrieval*, pp. 1-10, Las Vegas, NV, March 1992 (invited).

[GN95] G. Nagy, "Document Image Analysis: Automated Performance Evaluation," *Document Analysis Systems*, A. L. Spitz and A. Dengel, Eds., pp. 137-156, World Scientific, Singapore, 1995.

[GT90] "Counseling Women in Engineering Technology to Prepare for Their Future," S.A.R. Garrod and M.R. Taber, *1990 ASEE Annual Conference Proceedings*, pp. 784-788, 1990.

[H05] "The Black Box Report: Critical Security Issues with Diebold Optical Scan Design," Harri Hursti, Black Box Voting Project Report, July 4, 2005.

[H06] "Diebold TSx Evaluation: Critical Security Issues with Diebold TSx," Harri Hursti, Black Box Voting Project Report, May 11, 2006.

[I00] Idaho 2000 Census Results. http://www2.census.gov/census_2000/datasets/demographic_profile/Idaho/2kh16.pdf

[KMRS06] "Security Assessment of the Diebold Optical Scan Voting Terminal," A. Kiayias, L. Michel, A. Russell, and A. A. Shvartsman, UConn Voting Technology Research Center, October 30, 2006.

[KSRW04] "Analysis of an Electronic Voting System," Tadayoshi Kohno, Adam Stubblefield, Aviel D. Rubin, and Dan S. Wallach, *IEEE Symposium on Security and Privacy*, May 2004.

[L06] "Three Computationally Demanding Problems in Search of ASAP Solutions," D. Lopresti, *Proceedings of the Seventeenth IEEE International Conference on Application-specific Systems, Architectures and Processors*, September 2006, Steamboat Springs, CO (invited paper).

[LM06] "2006 Survey of Public Attitudes Toward Electronic Voting in Pennsylvania," Christopher Borick, Daniel Lopresti, and Ziad Munson, Lehigh University / Muhlenberg College Institute of Public Opinion, October 3, 2006, Department of Computer Science and Engineering Technical Report LU-CSE-06-35.

[LN05] D. Lopresti and G. Nagy, "Mobile Interactive Support System for Time-Critical Document Exploitation," *Symposium on Document Image Understanding*, College Park, MD, November 2005.

[LNT96] Y. Li, D. Lopresti, G. Nagy, and A. Tomkins, "Validation of Image Defect Models for Optical Character Recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 18, #2, pp. 99-108, February1996.

[LR05] "The Effectiveness of Generative Attacks on an Online Handwriting Biometric," D. Lopresti and J. Raim, *Proceedings of the Fifth International Conference on Audio- and Video-based Biometric Person Authentication*, T. Kanade and A. K. Jain and N. K. Ratha, eds., Berlin: Springer-Verlag, 2005, pp. 1090-1099.

[LTZZ95] "Systematic Bias in OCR Experiments," D. Lopresti, A. Tomkins, J. Zhou, and J. Zhou, *Proceedings of the IS&T/SPIE International Symposium on Electronic Imaging*, L. M. Vincent and H. S. Baird, eds., February 1995, San Jose, CA, pp. 196-204.

[LZ97] "Using Consensus Sequence Voting to Correct OCR Errors," D. Lopresti and J. Zhou, *Computer Vision and Image Understanding*, vol. 67, no. 1, pages 39-47, 1997.

[M05] "Electronic Voting," Rebecca Mercuri, http://www.notablesoftware.com/evote.htm.

[M06] "Paper Jams a Problem for Electronic Voting," Stephen Manning, *The Washington Post (Associated Press)*, December 21, 2006.

[MBS06] "Education by Design: Connecting Engineering and Elementary Education," R. Miller and E. H. Barney Smith, *Proceedings Hawaii International Conference on Education*, Honolulu, HI, January 2006, pp. 4539-4550.

[NE81] G. Nagy and D.W. Embley, "Behavioral Aspects of Text Editors," *ACM Computing Surveys*, vol. 13, #1, pp. 33-70, March 1981.

[NP06] "New York State Voter System User Rate Assessment Study," Dwayne G. Norris and Christine A. Paulsen, American Institutes for Research, December 11, 2006. http://www.elections.state.ny.us/NYSBOE/HAVA/DRAFTAIRSTUDY.pdf

[NRC91] "Women in Science and Engineering: Increasing Their Numbers in the 1990s." National Research Council, 1991.

[NRC92] "Science and Engineering Programs On Target for Women?" National Research Council, 1992.

[NS04] M. Norris and E. H. Barney Smith, "Printer Modeling for Document Imaging," Proc. of the International Conference on Imaging Science, Systems, and Technology: CISST'04, Las Vegas, NV, June 21-24, 2004, pp. 14-20.

[NSF96] "Women, Minorities, and persons with Disabilities in Science and Engineering: 1996." National Science Foundation, NSF 96-311, Sept. 1996.

[NSPE92] "The Glass Ceiling and Women in Engineering." National Society of Professional Engineers, #3054, Aug. 1992.

[P06] David Chaum's punchscan system http://www.punchscan.org/

[PA06] "Examination Reports Issued by the Secretary of the Commonwealth." http://www.hava.state.pa.us/hava/cwp/view.asp?a=1283&Q=445840

[PAV06] "Commonwealth of Pennsylvnia: Certification Videotape for Diebold Optiscan E-voting System," recorded November 22, 2005 in Harrisburg, PA (Michael Shamos, examiner).

[RB00] Robert B. Putnam, Bowling Alone: The Collapse and Revival of American Community, Simon & Schuster, 2000.

[RISC04] “Trusted Agent Report: Diebold AccuVote-TS Voting System,” RABA Innovative Solution Cell (Dr. Michael A. Wertheimer, Director), January 20, 2004.

[RNG99] S. Rice, G. Nagy, and T.A. Nartker, *Optical Character Recognition: An Illustrated Guide to the Frontier*, Kluwer Academic Publishers, Boston/Dordrecht/London, 200 pages, 1999.

[SA06] E. H. Barney Smith and T. Andersen, "Partitioning of the degradation space for OCR training," *Proc. SPIE Electronic Imaging, Document Recognition and Retrieval XIII*, Vol. 6067, San Jose, CA, January 2006, paper #606705.

[SLZN98] P. Sarkar, D. Lopresti, J. Zhou, and G. Nagy, "Spatial Sampling of Printed Patterns," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, #3, pp. 344-351, March 1998.

[SQ02] E. H. Barney Smith and X. Qiu, "Relating statistical image differences and degradation features,” *Document Analysis Symposium*, Princeton, NJ, 19-21 August 2002, pp. 1-12.

[SQ04] E. H. Barney Smith, and X. Qiu, "Statistical image differences, degradation features and character distance metrics," *International Journal of Document Analysis and Recognition*, Vol.6, No. 3, 2004, pp. 146-153.

[TGB98] “Retention and Performance of Male and Female Engineering Students: An Examination of Academic and Environmental Variables,” S. Takahira, D.J. Goodings, and J.P. Byrnes, *Journal of Engineering Education*, Vol. 87, pp. 297-304, July 1998.

[UD06] “Changes Are Expected in Voting by 2008 Election,” Ian Urbina and Christopher Drew, *The New York Times*, December 8, 2006.

[WHH04] “The Effectiveness and Trustworthiness of Faith-Based and Other Service Organizations: A Study of Recipients’ Perceptions,” Robert Wuthnow, Conrad Hackett, and Becky Yang Hsu, *Journal for the Scientific Study of Religion*, 43(1): 1-17.

[WHPTV06] “Parties outline recount proposals with House control at stake,” www.whptv.com (Associated Press), December 12, 2006.

[Y04] “Estimating Degradation Model Parameters from Character Images,” Hok Sum Yam, Masters Thesis, Boise State University, December 2004.

[YS03] H. S. Yam and E. H. Barney Smith, "Estimating Degradation Model Parameters from Character Images," *Proc. International Conference on Document Analysis and Recognition*, Edinburgh, Scotland, 3-6 August 2003, pp. 710-714

[ZL94] “Repeated Sampling to Improve Classifier Accuracy,” J. Zhou and D. Lopresti, *Proceedings of the IAPR Workshop on Machine Vision Applications*, December 1994, Kawasaki, Japan, pp. 346-351.

[ZLSN97] J.Y. Zhou, D. Lopresti, P. Sarkar, and G. Nagy, "Spatial Sampling Effects on Scanned 2-D Patterns," Advances in Visual Form Analysis, C. Arcelli, L. P. Cordella, and G. Sanniti di Baja, Eds., pp. 666-676, World Scientific, 1997.