The ElectionTracker Public Repository for Election Results: A New Tool for Greater Election Transparency

June 4, 2007

submitted in response to the solicitation, "Make Voting Work, Requests for Proposals: New Diagnostics and New Solutions," The PEW Charitable Trusts

Principal Investigator (PI) and Primary Contact
Dr. Alan T. Sherman
Associate Professor, Computer Science,
and Member, National Center for the Study of Elections (NCSE)
Dept. of Computer Science & Electrical Engineering (CSEE)
University of Maryland, Baltimore County (UMBC)
1000 Hilltop Circle Baltimore, MD 21250 USA
cell 410-963-4779, home 202-966-7204, office 410-455-2666
email: dralansherman@starpower.net, fax 410-455-3969

University Administrative Contact
Margarita Cardona
Senior Grants and Contracts Manager, Office of Sponsored Programs
ECS Building, Room 329
UMBC, 1000 Hilltop Circle Baltimore, Maryland USA
tele: 410-455-8635, fax: 410-455-1876, URL: www.umbc.edu

Summary Information
Type of Proposal: Planning grant
Proposed Period of Grant: September 1, 2007–August 31, 2008
Total Funds Requested: $199,996
UMBC Taxpayer Identification Number: 526002033, DUNS Number: 061364808
Keywords: Bulletin board of election results, database of election results, election results integrity, ElectionTracker, election security, election transparency, public policy, voting.

Primary Team Members
Alan T. Sherman (PI), Warren D. Smith (CoPI),
Nicholas Miller (Investigator), Konstantinos Kalpakis (Investigator),
Richard T. Carback III (Graduate RA)

The ElectionTracker Public Repository for Election Results: A New Tool for Greater Election Transparency

June 4, 2007

Abstract

We propose a pilot project to work with the Maryland State Board of Elections to create applications and an interactive website, *ElectionTracker*, as a new tool to support greater election transparency. ElectionTracker will enable election officials immediately and easily to post digitally signed unofficial and certified election data at voting machine and precinct levels—a granularity currently unavailable on any website. Unlike any currently available archives of election results, ElectionTracker's immediacy and detailed granularity will permit any concerned citizen to check their polling places and the web to assure themselves that election totals are being properly reported.

Our interdisciplinary team will design a national ElectionTracker system and prove its feasibility by implementing a demonstration version for Maryland. After testing this demo system in mock elections, we aim to field it in the 2008 Maryland elections. To ensure that the system is useful and easy to use, we will hold two user conferences at UMBC to solicit feedback on our initial design and on ElectionTracker's performance in our mock election. Our final report will include recommendations for national standards for reporting and storing election results.

Contents

1 Introduction

Transparency is at the core of public confidence, but that does not ring true in U.S. elections. Most voters don't know who won each race at their polling place even if that race was featured on the news. For voters to find out is highly inconvenient, expensive, and/or long-delayed, and the policies and procedures for doing so differ drastically in different states. Further, to check the *integrity* of the data you would want to compare the centrally-claimed totals with the underlying precinct- and machine-level data. That is even more difficult and often impossible: data are typically available only in the aggregate.

The best information Joe Voter can currently hope to find on the Internet is by parish or county, and posting that can take months and isn't in the news because just saying who won and by what percentage is more succinct. In Maryland, unofficial results are posted on the State Board of Elections website [9] on election evening, but without showing precinct or machine level results.

We find this situation unacceptable. In any modern democracy, anyone should be able to download off the Internet all precinct totals for all candidates in all races shortly after the election. These totals should be broken down by absentee ballots, provisional ballots, and ballots cast at each polling place machine. Voters should be able to look at the data from county to polling place to a specific optical scanner or other machine, and do various kinds of statistical computations and searches, with a few mouse clicks. Recount data should also be made readily available. Unfortunately, the U.S. is a long way from this ideal situation.

While all of this information is available to them, election officials lack the resources and tools to post all of it in an efficient fashion. Currently, the best tools available permit them to upload to a database and print out that database and a summary of it. These numbers can then be manually checked with printouts from counting machines or reports from polling places.

Our proposal is to design, test, and pilot *ElectionTracker*, a web and software application that will allow election information to be easily and securely posted to the Internet for convenient access by voters. It harnesses the power of online communications. If widely adopted, it should go a long way toward making U.S. elections transparent, verifiable, and fraud-proof.

2 Research Questions

At the end of election day in Maryland, each polling place prints the results of vote totals off voting machines. A precinct captain signs them, posts the signed printouts on a bulletin board for all to see, reports them to the local election board, which sends them to the State Board of Elections. Similar processes are carried out in most states. Our goal in this research is to put that bulletin board on the Internet, plus make elections officials "digitally sign" and date the data they post in such a way that we get certain security guarantees. We want this to be efficient, uniformizing, robust, and secure. This system is something that could be used by all states, regardless of differences in their voting system technologies and their results posting processes.

Our pilot study will address: (1) can this be done?, (2) how (and what are the requirements)?, and (3) how well?

ElectionTracker will be useful in several ways. Most importantly, it provides voters and observers a complete view of the election. Each machine total, polling place total, and county/parish total could be checked to make sure that reported totals at each level match. Anomalies, such as the ones we discuss in §2.2-2.3 would be much more likely to be spotted.

ElectionTracker is inarguably more convenient for interested voters and elections observers and analysts. They can vote, and then at the end of the day go online and check vote totals for their polling place and all the way up the chain. That ability might be interesting to voters in and of itself (*e.g,* "The news says that my county voted for Bush, but how did *my* neighborhood vote?").

ElectionTracker will let elections officials post results efficiently online without having to develop their own system in-house. Implementing a system like ElectionTracker is nontrivial because there are some tricky security issues involved, and also because of the horribly-nonstandardized, non-specified, and non-uniform state of election equipment and data formats in the USA—which is another unbelievable disaster on which we'd like to have a positive impact. Our work will also serve as a guide for how to develop similar systems in the future.

Providing data at the machine-level granularity and digitally signing it can help increase the security of elections. The digital signature ensures that data are verifiably from the declared source and have not been altered since it was first written. The machine level granularity can add the ability of researchers to do many powerful statistical tests on the data, and investigators can use that information to pinpoint where to begin their investigations. The system can also provide a full "revision history" (with each revision timestamped and signed) yielding increased accountability and greater difficulty of fraud.

2.1 The usual reassuring story ...

The usual story is that US elections are secure because there are precincts and it is difficult and risky to organize a huge conspiracy to have many precincts fake their data.

If Central fakes the data from some precincts, then that alteration is detectable because all the precincts publish their data, and so does central, and somebody would scream after detecting the discrepancies between these published documents.

In summary: Central can't fake it (would be detected), and the precincts can't fake it (too big and risky a conspiracy), so we're safe. While this is a great story, the trouble is there is little relation between it and reality. That's a big reason we need to do something about it. The research question then is: what?

2.2 ...Versus reality I: Gundlach and Alabama

Auburn University Professor James Gundlach's statistical study [6] suggested fraud in Baldwin County threw the 2002 Alabama governor election from Siegelman(D) to Riley(R). Appendix B tells more of that story including sketching the inculpating statistical techniques. A planned deeper study by Gundlach jointly with Smith (WDS) was to examine more and better data from the 2006 elections using more powerful statistical methods, and Alabama agreed to give us the data (official precinct vote counts).

That was agreed before the November 2006 election. As of June 2007 and Alabama had still not provided any data. The moral is that the "usual story" above has little real validity. No external verifier ever checked the official versus actual precinct counts in the Alabama 2006 elections because those data simply were not available for the next 6 months.

That's not to say that Alabama is especially bad. Different US states have wildly different policies about releasing such election data, and Alabama is considered by (the USA's few and extremely underfunded and undermanned) election protection organizations to be one of the better ones.

We are not aware of *any* state that releases precinct counts immediately on the Internet, preventing almost the *only* way that statistics could be used to cast suspicion on an election soon enough for anybody to try to do anything about it.

2.3 ...Versus reality II: Gore's negative total in Florida.

At 2am, Al Gore suddenly received *negative* 16022 votes from precinct 216 in Volusia County Florida during the 2000 Presidential election. This was supposedly due to a faulty memory card and happened despite parity check bits on the memory card designed to make the probability of undetected random data alteration be at most $1/65536$. Also, amazingly, the same memory card delivered about 4000 erroneous Bush votes and 10000 votes for the Socialist Worker's Party candidate (approximately equal to his total in the entire rest of the USA), which positive votes *exactly* canceled Gore's negative votes so that the total vote number came out correct. This was later shown by a CBS internal investigation to be the direct cause of television networks calling Florida for Bush, which in turn was the direct cause of Gore's concession (which Gore later retracted, and a later upload from a different memory card repaired these particular data).

Regardless of whether you accept this official explanation (or how outrageous it is for any vote tabulator to silently accept "negative votes"), what matters for the present purposes is simply that if all precinct totals were instantly published on the Internet, then the chance that such anomalies would remain undetected for long, would be zero. As it was, this error was detected only by pure luck by an alert Democratic poll watcher who noticed Gore's total *decrease.*

2.4 What we propose to do about it

We propose creating a software system to allow instant publication of election results on the Internet. It will be both a pilot study of how hard it is to do this and how to do it, and (hopefully) proof that it can be done, and (hopefully) some attempt to test it to evaluate how well it works.

Once this system is available and proven to work, it would then be possible to hope for, e.g, legislation to mandate its use. Then and only then will our democracy be capable of becoming transparent, verifiable, convenient, secure, and efficient—at least at the (small) level described in the "usual story."

ElectionTracker won't deal directly with the promising emerging "end-to-end (e2e) secure" voting technologies (e.g. Punchscan [12, 4], 3Ballot [13], and VoteHere `www.votehere.com`), and won't be about alternative systems for collective decision making such as "approval" [2] and "range voting" [3]. However, it will help pave the way for them in the sense that, if ElectionTracker became widely used, it would become easier to switch to these revolutionary technologies.

2.5 Previous Research

There have been previous databases of election data placed on the Internet either for research or commercial purposes (*e.g.*, NEDA, ROADS, ANES, ICPSR) [11] or (to a limited extent) by governments. Check "secretary of state" and/or "board of elections" and "federal election commission" websites at state and local governments; these often post election results. Maryland, in particular, is very good about posting elections results as soon as is possible, which makes it an ideal candidate for this pilot system, because they have already allocated resources to Internet posting.

However, to our knowledge none of them use the powerful security tools that are available that we intend to use. Also, none post contemporaneous (as opposed to historical) data at the granularity of precinct and machine.

So we are entering new territory as far as Internet election data are concerned. On the other hand, we are *not* entering new territory concerning humans securely modifying important data stored at remote locations, and/or publishing some such data. Bulletin boards and blogs pervade the Internet. You can currently perform operations to your bank account and do stock trading online. Evidently, these problems have been, for practical commercial purposes, solved.

3 Approach

We now describe our approach by considering system architecture, security, database design, location, convenience, accuracy, evaluation, and plans for the full national system.

3.1 System Architecture

ElectionTracker will consist of:

1. a central hub, or "server," (a web site hosted on a computer in a secure location and communicating with the outside world only though a "firewall" second computer to prevent it from being "hacked") that stores the data, makes it publicly available and accepts updates, and backs up the data to protect against hardware failures.

2. "Viewers." Viewers can be any computer anywhere on the Internet running any standard web browser. They read data posted by the server. Ordinary people could also be allowed to post comments, suggestions, questions, and answers about system usage, but this forum will emphasize that it is *not* an official complaint site for election incident reports.

3. "Clients." Clients can be any computer anywhere on the Internet.[1]. They run special "client software" (downloadable in public-source digitally-signed form by anybody from the server). This client software's purpose is to allow a person who has it, to communicate with the server (or in some cases other

[1] Maryland "local boards of elections" have computers with internet and can scan paper "poll tapes." We will encourage precinct captains to verify results ASAP.

clients e.g, via email) using the cryptographic protocols we apply. Clients can then post or update election data, because their client software will take care of the digital signing, key creation + authentication + posting, time stamping, cryptographic, or whatever protocols. However, they will only be able to do this if they have the right keys. Some Random Joe can download our client software but will not have the keys and authorizations he needs to use it in evil ways. He will be able to use it in non-evil ways, though, e.g. to verify or deny the validity of digital signatures and timestamps.

See appendix C for cryptology terms like "time stamping," "keys,", and "digital signature."

The point of all this is not only to prevent Random Evil Hackers and malware from corrupting the data, but also to create confidence in the data's integrity and authorship (who and when) in spite of the possibility that perhaps we (the owners of the server) or corrupt government election talliers, are trying to fake the data.

3.2 Security

ElectionTracker will protect the availability, integrity, and authenticity of its data and services through a combination of standard security technology, procedures, and training of its users. Security technology will include physical security of the servers, key generation and distribution and management including a public-key infrastructure [8], digital signatures, cryptographic hash functions, and (possibly) physical tokens to authenticate official users. All of this technology (while nontrivial) is well developed and our team is well-qualified to apply it properly. Most importantly, anyone can check and attempt to verify any ElectionTracker data.

Let's go though the most obvious possibilities.

1. Evil server owners or external attackers (or just, e.g, non-evil power outages) could destroy the server. But this is somewhat preventable by having many "mirror" servers and via firewalls and secure locations. We shall want two, just to prove we can do mirroring.

2. Evil server owners could fake the data on the server. But this is computationally infeasible to do without immediate external detection because of invalid timestamps or digital signatures.

3. Evil server owners could compromise the client software in secret evil ways. But that software is public-source, not secret. Hence any such gambit would risk exposure at any later time by any later inspector of the code, whereupon they'd land in jail.

4. Evildoers could distribute bogus "spoofed" client software. But this software would necessarily fail to have our correct digital signature and hence would be instantly detectable (plus the correct software is downloadable from the server for comparison).

5. Evil election officials could try to change the data in evil ways. But the server records (and publicizes) the entire update history of all data, and who did each update and when. And the who/when/what information is unalterable and undeleteable without immediate detectability by anyone in the world.

6. Evil election officials or other interested Evil parties could coerce or bribe non-evil election officials to give them their secret keys. From then on, the evil ones would be able to masquerade as the non-evil ones and do whatever they are allowed to do. However, we could make revelation by an election official of his secret key (to anyone, including to his bosses or us, whether intentionally or not) be a crime and job firing offense. Further, the server could record secret information about who did what when from where on the Internet, which could possibly be used to prove some election official must have given away his key (you weren't in Kalamazoo Michigan at that time, but somebody with your key was!). That possibility makes any revelation of keys very dangerous but remains the plan's biggest weakness.[2]

Some people might wonder if ElectionTracker might unduly compromise ballot privacy in the case that very few people vote on a particular machine. The theoretically possible worst case is only one person voting on a machine. We reply that machine-level tallies are public information in Maryland (albeit not easily available), so ElectionTracker would not greatly change the status quo. Second, although Maryland records who

[2] It may be possible to add some additional traps for key-stealers that would create risks that those using such keys might actually merely be creating incontrovertible evidence usable against them in court. Other technical strategies include protecting keys with physical security and splitting keys into shares.

votes on each machine, that information is not public. Third, typically 50-200 people vote on each machine in Maryland and machines are unlikely to have very few voters. The core problem is not ElectionTracker but limitations of current voting technology; it is sound government policy to post machine-level tallies.

3.3 Database Design

ElectionTracker will be based on modern database web applications technologies [5], which offer numerous advantages in storing, retrieving, and managing large volumes of data efficiently. They provide mechanisms for keeping the data consistent across multiple "mirror" duplicate databases despite concurrent manipulation of such information by many users and recovery of data in case of system failures. We envision both a database server and a web application server. Most of it will be built from commercial off the shelf software and hardware.

We will develop simple intuitive dynamic web interfaces for voters to access ElectionTracker. The system will be mirrored for increased reliability with all servers located safely behind a firewall.

3.4 Location

The pilot project and testing for ElectionTracker will be done in Maryland. We hope to have one full county participating for the November 2008 election.

Implementation and design will be completed at UMBC. UMBC, with its robust Internet services, will also host all Internet servers and equipment we will use during the pilot project in a secured computing room, completely separately from all other university computer systems.

3.5 Convenience

The success of ElectionTracker will depend crucially on the the system being easy to use, both for officials and for voters. This (and security) will be our highest priorities. We will publish user guides, help pages, promotional and educational material, and design specs, on the server itself, perhaps with interactive demonstrations. These all should make user-training easier than in the pre-computer era. We will also make a video for these purposes.

At least initially, ElectionTracker will be unofficial and in addition to the current procedures in place in Maryland for aggregating votes. As such, it will require additional work, including installing software, training users, generating and managing keys. However, in Maryland, each local election board uploads unofficial data to the media on election evening, and having that official also upload the same data to ElectionTracker would be a simple task. Officials will not have to change the data formats they use. ElectionTracker will perform any required data format conversions.

Ultimately, we envision that using ElectionTracker will be *easier* than not using it, because it provides essential services at higher levels of security and uniformity. Eventually, election officials might choose to replace some existing official procedures with ElectionTracker.

3.6 Accuracy

With the ease of data updates ElectionTracker provides, and its uniformity of formats and automation, and the ease of checking by many more people, accuracy should increase.

3.7 Evaluation

We aim for the goals of greater convenience, uniformity, accuracy, security, and transparency than now, and we want to evaluate how well we've met those goals. By building public question/comment and user-polling features into the system, such evaluation and feedback will happen almost automatically (perhaps with a little encouragement needed from us), and continuously and perpetually, as a natural side effect of use and of interest in the system from the crypto, software, and election communities. We will hold two user conferences to listen to concerns of users; survey users; and hold focus groups. We will invite informal "Red Teams" to probe our designs and implementations for possible security weaknesses, within the scope possible of our resources. E.g. this could be an assignment in an information assurance course.

4 Scope

Our vision is for a national ElectionTracker system. Toward that vision, in this one-year pilot, we will focus on a limited demonstration system in Maryland. To make best use of limited time and resources, we it's wisest to restrict the demonstration system as much as reasonably possible, while not trivializing the challenge. Follow-up projects could then increase the scope of the effort, within Maryland, and beyond.

To this end, we propose a demonstration system at the county level (*e.g,* Howard County, near UMBC). By limiting the number of precincts involved, and by limiting the geographic region of the demonstration (Maryland stretches from ocean to mountains), the effort will be more manageable. Following the Maryland test, we hope the project could ultimately scale up to national level (all 50 states).

In Maryland, there 23 counties plus Baltimore City, 1,789 precincts, approximately 20,000 Diebold AccuvoteTS electronic voting machines, and one central scanner. There are about 3.1 million registered voters (out of 5.6 million residents) who vote on 70 elected offices. If Senate Bill 392 is funded, then Maryland will eventually switch to an optical scan system, purchasing one scanner and one ballot marking device per polling place.

Limiting the scope of the project also involves choices of level of service (fanciness of user interface, functionality of statistical analysis, level of security provided, resilience to failures, number of kinds of format conversions supported, choices in key management and delivery, etc.). "Red Team" security analyses are very useful but expensive.

Given the difficulties of software development including cost estimation, we feel it is best to err on the side of conservative choices with regard to scope in this first year. A demonstration at the county level is complex enough to reveal most of the issues confronting a national system.[3]

5 Work Plan / Timeline for major tasks

Sept. 1, 2007, Begin project
November 1, 2007, Finish preliminary design document for demo system (2 months)
Early November, 2007, Host User Conference I
December 15, 2007, Finish design document for demonstration system (1 month)
April 15, 2008, Finish preliminary implemention of demo system (5 months)
June 15, 2008, Finish testing, including trial of demo system in mock election (2 months)
Late June, 2008, Host User Conference II
August 15, 2008, Finish evaluation of demonstration system and final report (1 month)
August 31, 2008, Year 1 funding ends
November 2008, Use demonstration system in Maryland Election

6 Staffing Plan

Our interdisciplinary team includes experts in the key areas of database systems, security, and web programming, as well as voting theory and election processes. The project will also benefit from interactions with other experts at UMBC.

• **Alan T. Sherman (PI)** (PhD, Computer Science, MIT, 1987). Expert in information assurance, cryptology, and security of voting systems. Associate Professor of Computer Science, Director of the DoD-funded UMBC Center for Information Security and Assurance (CISA), member of the National Center for the Study of Elections at UMBC, an editor of *Cryptologia*, and Organizer of the NSF-funded VoComp University Voting Systems Competition. Research includes high-integrity voting systems, including Punchscan, and key management. As a member of the Maryland Study on Vote Verification Technologies [15], and as an expert witness on voting system security, Sherman has examined in detail voting procedures used in Maryland. Sherman is one of the few people in the world who teaches a course on electronic voting systems. `www.umbc.edu/engineering/csee/faculty/sherman.html`

[3] "Early publication" issues will be more severe if we go national; we propose ElectionTracker not publish results until polls close.

Sherman will act as Principal Investigator, overseeing interactions with SBE and other investigators. He will be involved in all aspects of the project, and ensure that the project is on track for the 2008 election.

- **Warren D. Smith (Co-PI)** (PhD, Applied Math, Princeton, 1988). Co-founder of `RangeVoting.org` voting reform and educational organization and creator of 90% of its web content and programming. Creator of the ElectionTracker concept. Author of numerous scientific papers, including many on voting and cryptology-related topics; recent papers online at `math.temple.edu/∼wds/homepage/works.html`. Will be involved in all levels of the system design and implementation, especially security issues.
- **Nicholas R. Miller (Investigator)** (PhD, Political Science, UC Berkeley, 1973). Arrived at UMBC as Instructor in 1971; now Professor of Political Science. His principal research lies in the area of formal political theory dealing with collective decision making and, in particular, formal theories of voting processes. His research on logrolling, majority voting, power, social choice, information pooling, agenda control, and spatial voting models, has appeared in such journals as Amer. Political Science Review, Amer. J. Political Science, Political Analysis, Public Choice, Mathematical & Computer Modelling, Theory & Decision, and J. Theoretical Politics, in a number of edited books, and in a monograph on Committees, Agendas, and Voting. Six years as receiving editor of the J. of Theoretical Politics, now eJournal editor for Social & Public Choice Theory in the new Political Science Network of the Social Sciences Research Network. `www.research.umbc.edu/∼nmiller/`

Miller will play a key role in the requirements, design, and evaluation stages of the project, providing insight into the voting processes and how they change from state to state.

- **Konstantinos Kalpakis (Investigator)** (PhD, Computer Science, UMBC, 1994). Associate Professor of Computer Science, UMBC. Kalpakis' research focuses on content-based indexing and retrieval for multimedia databases and investigating issues in feature extraction, similarity metrics, classification, indexing, and replication. Also designing, in cooperation with NASA/CESDIS and the Law Library of the Library of Congress, the architecture of an international legal digital library with over 12 member countries. `www.umbc.edu/engineering/csee/faculty/kalpakis.html`

Kalpakis provides expertise in database design and data manipulation. He will work on the database schema, data conversion, reliability, and disaster recovery.

- **Richard T. Carback III (Graduate Assistant)** (BS, Computer Science, UMBC, 2005). MS/PhD student in computer science working in Information Assurance at UMBC. Currently, he is working on the Punchscan e2e-secure voting system and assisting with VoComp (`www.vocomp.org`, an intercollegiate voting systems competition fostering innovations in voting system technology) & serving as a website programmer for both. Over 5 years experience developing enterprise level database and web applications for L3 Communications, Government Services, Inc. `www.cs.umbc.edu/∼carback1/`

Carback's web programming expertise is well matched for this project and he will be responsible for application development. He will supervise hourly programmers and work with the investigators to ensure that the system meets all of its specifications.

- **Hourly Programmers** will be hired for the project throughout the year to work on application development. We plan to hire two upper-level computer science undergraduates for programming, and a graphic design undergraduate for image manipulation and design assistance.
- **National Center for the Study of Elections (NCSE) at UMBC** — see Appendix D

7 Budget and Budget Justification

Direct Costs (2007-2008)

Salaries	**Note**	**Rate**	**Number**	**Amount**	**Subtotal**
Alan Sherman (PI)	summer support			10,942	
Nicholas Miller (Investigator)	salary support			10,942	
Warren Smith (CoPI)	salary support			45,000	
Kostas Kalpakis (Investigator)	summer support			10,942	
Richard Carback (grad assistant)	12-month RA			23,000	
additional programmers	hourly	20	1,000	20,000	
subtotal:					**120,826**
Benefits					
Sherman - fringe benefits		30%		3,283	
Miller - fringe benefits		30%		3,283	
Smith - FICA, unemployment tax		9%		4,050	
Kalpakis - fringe benefits		30%		3,283	
Carback - health benefits				1,871	
Carback - tuition	per credit hour	441	20	8,820	
subtotal:					**24,589**
Other					
Equipment				12,000	
Documentation Costs: equipment, supplies, travel				6,600	
Software Licenses				3,800	
Conference Travel		2,000	4	8,000	
User Conferences		3,000	2	6,000	
subtotal:					**36,400**
Total Direct Costs					**181,815**
Indirect Costs					
UMBC overhead on direct expenses		10%		18,181	
subtotal:					**18,181**
Total Costs					**199,996**

1. **Salaries.** Expertise is needed in key areas of government policy, security, database systems, web programming, voting theory & election processes, and project administration. Each UMBC faculty team member will receive a flat $10,942 summer or salary support for his time. The translation of this amount into percent effort varies with the individual and his 9-month salary. At UMBC, one summer month of salary is 1/9 of the 9-month salary. Release time from teaching is computed as 1/8 of the 9-month salary, per semester course reduced, should any faculty choose to apply their salary support in this fashion. Additional programmers for demo system will be paid hourly at $20 per hour. All team members are at UMBC, except for Warren Smith, who will devote approximately 50% of his effort to the project. If funded, we expect to hire Smith at UMBC as a visiting research scientist in CSEE.

2. **Benefits.** UMBC charges 30% fringe benefits on all salaries, except for graduate assistants and hourly workers. Smith declines benefits; UMBC charges 9% for FICA and unemployment tax. Graduate assistants receive tuition and health benefits ($1871). Carback will be a full-time, 12-month graduate assistant in the Dept. of CSEE. His tuition benefits are calculated assuming 20 credits per year at $441 per credit hour.

3. **Equipment.** Equipment is needed for the demo system, including an application server ($4k), database server ($4k), backup system ($1.5k), firewall, switches, cables, physical authentication tokens for election officials, and lab supplies ($2500).

4. **Software licenses.** For Adobe CS Mac 2x ($399 each), SSL Certificates (2 years - $3000).

5. **Travel.** To present results at conferences and to interact with election experts and officials. Confer-

ences might include NASS, Usenix/Accurate EVT. Expenses will include registration, travel, lodging, subsistence.

6. **User conferences.** To seek input and feedback from users of ElectionTracker, including for the demo system, we will host two user conferences at UMBC. One will be held after our initial design; the other will be held after a mock election trial of the demo system. Expenses will include room reservations, room setup, AV services, information materials given to participants, advertising, subsistence.
7. **Documentation costs.** To make on-line documents, manuals, and promotional materials to train and inspire users. Equipment includes digital video camera and software: 2x HD video cameras ($1000 each), microphones and tripod ($300), fast 200 GB disks for editing ($800), Final Cut Pro editing software ($1000), DVDs, videotape, travel for documentation video ($2500).

Note on cost sharing: Charles Nicholas, Chair of UMBC's CSEE Dept, has promised the following cost-sharing incentive. If this proposal is funded, he will reduce the teaching load for Sherman or Kalpakis by one semester course per year, without cost to this project, to make more time available to carry out the work.

Acknowledgments. We appreciate helpful comments from Tim Brennan & Don Norris (Public Policy), Tim Finin & John Pinkston (CSEE), and Ross Goldstein (Maryland State Board of Elections).

References

[1] I.F.Blake, G.Seroussi, N.P.Smart: Elliptic Curves in Cryptography, London Math'l Society Lecture Notes #265, Cambridge University Press 1999. Errata www.hpl.hp.com/infotheory/errata082900.pdf. Includes some elliptic curve cryptography/signature standards documents as appendices.

[2] S. Brams & P. Fishburn: Approval Voting, Birkhauser, 1983.

[3] Center for Range Voting http://RangeVoting.org/.

[4] Kevin Fisher, Richard Carback, and Alan Sherman: Punchscan: Introduction and System Definition of a High-Integrity Election System. In Peter A. Ryan, editor, *Pre-Proceedings of the IAVoSS Workshop On Trustworthy Elections (WOTE'06),* Cambridge, UK, June 2006.

[5] H. Garcia-Molina, J.D. Ullman, J.D. Widom: Database Systems Complete Book, Prentice Hall, 2002.

[6] James H. Gundlach: A Statistical Analysis of Possible Electronic Ballot Box Stuffing: The Case of Baldwin County Alabama Governors Race in 2002, Presented at the Annual Meeting of the Alabama Political Science Association, Troy Alabama, 11 April 2003. http://www.auburn.edu/∼gundljh/Baldwin.pdf.

Update: Election Theft, A Second Look: The Case of Baldwin County in Electing Alabamas Governor in 2002, APSA Meeting 1-2 April 2005 Jacksonville State Univ. http://www.auburn.edu/∼gundljh/Baldwin2.pdf.

[7] S. Haber & W.S. Stornetta: How to time-stamp a digital document, J.Cryptology 3,2 (1991) 99-111.

[8] R. Housley, W. Polk, W. Ford, and D. Solo, "Internet X.509 Public Key Infrastructure Certificate and Certificate Revocation List (CRL) Profile," RFC3280 (2002).

[9] Maryland State Board of Elections web site. http://www.elections.state.md.us

[10] R.C. Merkle: Protocols for public key cryptosystems, IEEE Sympos. Security & Privacy (1980) 122-134.

[11] Some important election datasets:
NEDA (Nat'l Election Data Archives) http://uscountvotes.org, a database that you can contribute to, but has largely been unused; ROAD (Record of American Democracy) http://data.fas.harvard.edu/ROAD/, unusual in that it includes precinct-granularity data, but it and the following all are historical not contemporaneous; ANES (American National Election Studies) http://www.electionstudies.org/ dataset at U.Michigan, heavily used by political scientists but their data are entirely survey-based; ICPSR (Interuniversity Consortium for Political and Social Research) at U.Michigan http://www.icpsr.umich.edu/ is the primary social science data archive in the U.S. and perhaps world, mainly a repository for survey data collected by social scientists but also holds some historical election and government data. Here's a dataset of elections made available for researchers by one of us (gives *all ballots* in every election) conducted with more-advanced voting systems than plain plurality voting http://RangeVoting.org/TidemanData.html; finally Election Data Services is a company with database of US presidential, congress, senate, governor races with county-level results since the 1940s, it currently charges $950 per data-year for access.

[12] Punchscan website. http://www.punchscan.org

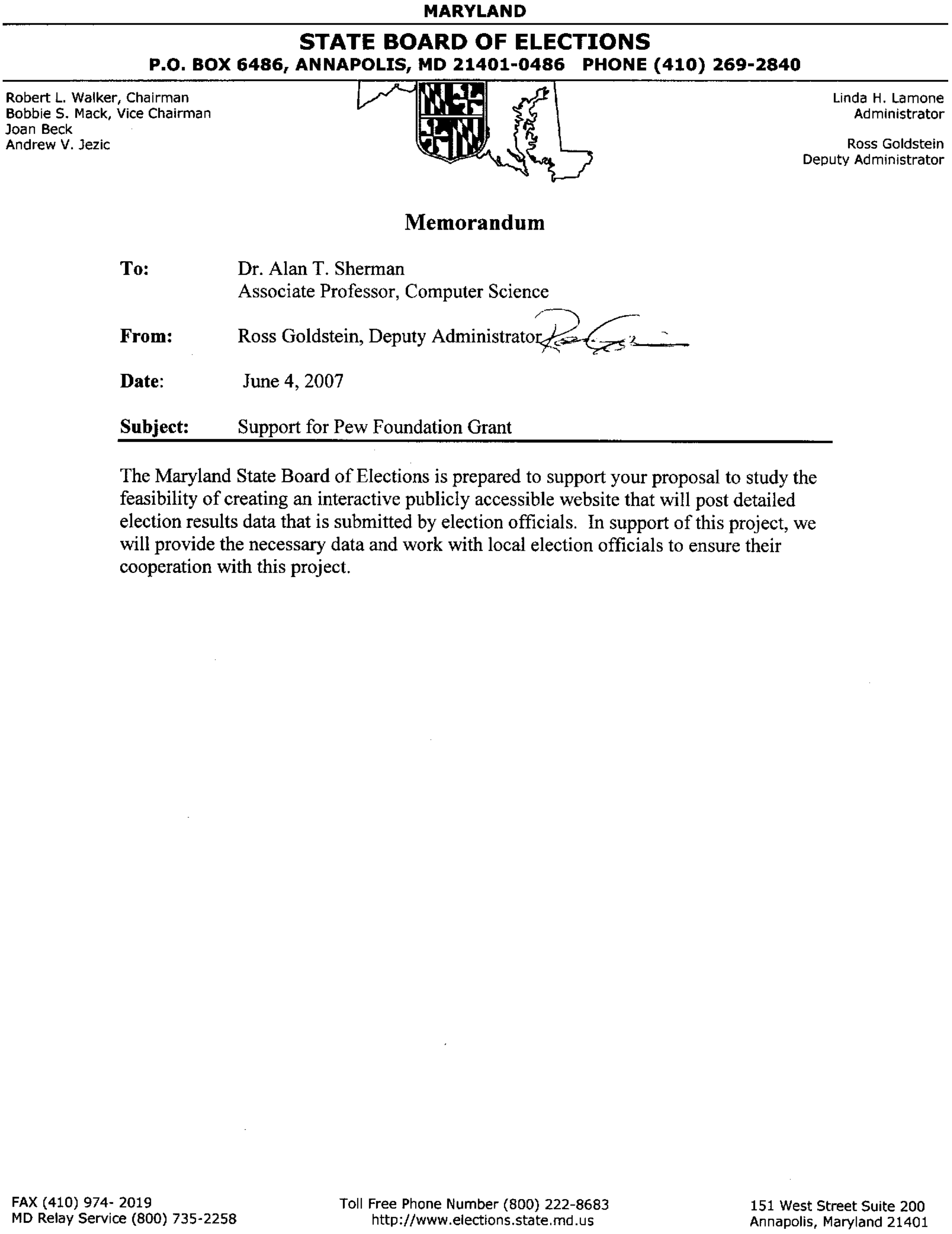
[13] R. L. Rivest & Warren D. Smith: Three Voting Protocols: ThreeBallot, VAV, and Twin, paper #99 at
http://www.math.temple.edu/∼wds/homepage/works.html, accepted EVT07 Boston Voting Conf.

[14] Aviel D. Rubin: Brave New Ballot, Morgan Road 2006.

[15] Alan T. Sherman & 9 others: An examination of vote verification technologies: Findings and experiences from the Maryland Study, *Proceedings of the USENIX/Accurate Electronic Voting Technology (EVT) Workshop* (August 2006). http://www.usenix.org/events/evt06/tech/

[16] Warren D. Smith: Cryptography meets voting, paper #80 at
http://www.math.temple.edu/∼wds/homepage/works.html.

A Letter of support from Maryland State Board of Elections



MARYLAND

STATE BOARD OF ELECTIONS

P.O. BOX 6486, ANNAPOLIS, MD 21401-0486 PHONE (410) 269-2840

Robert L. Walker, Chairman
Bobbie S. Mack, Vice Chairman
Joan Beck
Andrew V. Jezic

Linda H. Lamone
Administrator

Ross Goldstein
Deputy Administrator

Memorandum

To:	Dr. Alan T. Sherman Associate Professor, Computer Science
From:	Ross Goldstein, Deputy Administrator
Date:	June 4, 2007
Subject:	Support for Pew Foundation Grant

The Maryland State Board of Elections is prepared to support your proposal to study the feasibility of creating an interactive publicly accessible website that will post detailed election results data that is submitted by election officials. In support of this project, we will provide the necessary data and work with local election officials to ensure their cooperation with this project.

FAX (410) 974- 2019
MD Relay Service (800) 735-2258

Toll Free Phone Number (800) 222-8683
http://www.elections.state.md.us

151 West Street Suite 200
Annapolis, Maryland 21401

B More details about Alabama 2002 Governor election

We here provide more details than §2.2, extracted from numerous press accounts, Gundlach's papers, and Alabama statutes, all of which can be provided on request. The initial count showed Siegelman won. Then in Baldwin County – one of the few where all of the county officers in charge of elections were Republicans – the probate judge for elections (a Republican) corrected the tabulation, after midnight and long after poll watchers and other staff had gone, due to his conclusion that Siegelman had 6334 too many votes. Hence Riley won the governorship.

(The alteration in the absence of watchers violated Alabama law.) County officials were vague about the cause of supposed error ("lightning strike," "computer glitch").

Gundlach [6] speculated that the "correction of the error" actually consisted, basically, of simply removing about 20% of Siegelman's votes and awarding them instead to Riley, in each Baldwin precinct. This fraud, if such it was, could presumably have been detected by examining the "poll tapes" from every voting machine in the county to see what the original (as opposed to altered) data were. But attempts by Siegelman's Democrats to have such an examination failed, because Alabama's attorney general William H. Pryor took it to the Alabama supreme court to insist on the state's right not to have that data ever revealed and not to have any recount done with Democratic party or independent observers. He won the case.

All the judges (and Pryor) were either elected Republicans or appointed by Republicans.[4] With these data now permanently sealed, what evidence was there for a fraud? Gundlach employed the following statistical idea. Predict the 2002 Riley and Siegelman votes in Alabama precincts (both in Baldwin and in other counties) from corresponding counts in other races and/or in other years. These predictions exhibit 96-98% correlations with the official vote counts in both Baldwin and the other counties *but*, amazingly, all of the Baldwin precincts in 2002 showed a remarkably uniform percentage drop of Siegelman votes below the predictions. This happened *only* for Baldwin precincts and not for precincts in other counties.

ElectionTracker could make many such kinds of "forensic statistical analysis" easy and rapid (*e.g.*, 1 day). In contrast, Gundlach's final analysis still has not been completed over 4 years later.

C Cryptology terms of the art

We cannot describe cryptographic methods precisely here [1, 16]; we merely want to give the reader at least a vague notion of the key cryptographic tools for us and their purpose. They are:

1. digital signatures and public key cryptography,
2. secure time stamping,
3. hierarchy techniques.

Public key cryptography refers to the notion that each person could have a publicly known key which anybody could use to send a secret message to that key's owner using a standard published encryption algorithm. The encrypted message would only be readable by the key's owner and not anybody else. (By "key" we mean "a particular bit sequence"; by "publicly known" we mean, Joe's key is published in a directory under Joe's name and Joe publicly agreed he created it.)

If Joe sends you a *digitally signed* message, then you can be confident – and this can be regarded as proof valid in court [5] – that message came from Joe (or somebody who knew Joe's secret "key"), because it is computationally tremendously infeasible for anybody else to forge bits representing a correctly signed message. (The forgery work to guess somebody's key or successfully act as though you have can easily be made to far exceed what could be done by say, 1 computer per atom in the known universe, working for the entire age of the universe. The mathematical proofs of this and other security claims demonstrate that anybody who could break the security guarantees would be able to solve certain standard mathematical

[4] Alabama law was later changed so that, supposedly, with the present law Pryor would not have won the case. However, this law change, being *ex post facto*, left the 2002 election unaffected.

[5] Digital signature schemes have been standardized both by the US government and international standards agencies and carry the force of law in the USA under the Millenium Digital Commerce Act of 2000.

problems, such as the "discrete logarithm problem in elliptic curve groups" that are believed to be very hard and whose difficulty is continuously gaugeable by perpetually ongoing competitions with large money prizes to solve such problems of different sizes.) It is, however, computationally easy for *anybody* who has the signed message and knows Joe's *public* key to verify (or deny) the validity of the signature and to verify that it is Joe's signature and not anybody else's.

With *secure time stamping* [7], people post messages to a publically readable bulletin board, which are then time-stamped. Anybody looking at the information posted on the bulletin board can be confident that the messages arrived in the order claimed by the time stamps. That is because forging fake messages, or faking the time stamp bits, with the goal of inserting or deleting messages or changing the claimed time-order, is, again, hugely computationally infeasible. (The time stamps are bit sequences that depend on the previous messages and stamps in ways exceedingly hard to fake. One can also do a different kind of time stamping just by including the time in the messages, and having all parties digitally sign them to say, in an unforgeable and undeniable way, that they agree with the claimed time.)[6] Again, Joe Message Poster cannot hope to deny, in a court of law, that he posted it at time X.

One can also make hierarchies (trees) of public keys for use in digital signing or public key cryptography. The point of these techniques for us, is that they allow, e.g. the State Election Commissioner to be the root of the tree and authorize the County Election Commissioners (next level down in the hierarchy). They in turn can authorize Town Heads (next level down) who in turn can authorize Precinct Heads. We then have a tree of authorizations allowing various election officials to create their public and secret keys and to use them to post or update signed-by-them and timestamped data on our bulletin board, *but* only the data they are pre-authorized for, and not other data [10]. Each "boss" verifies (and digitally signs) proofs that his "underlings" one level down indeed created their keys (and not some random impostor).

Once you appreciate these technologies, it then is not hard to conceive of ways to use them to create a reasonable ElectionTracker system.

D About UMBC and NCSE

UMBC is a mid-sized public university located in the Baltimore-DC corridor. The Carnegie Foundation ranks UMBC in the category of Research Universities with high research activity. Students majoring in information technology and in visual arts make up a significant fraction of UMBC's 12,000 students – approximately 2,000 majors last year. UMBC is a DoD/DHS-designated "Center of Excellence for Information Assurance (CAE)" thanks to its *Center for Information Security & Assurance (CISA)*.

UMBC's Meyerhoff Scholarship Program (founded 1988) is dedicated to increasing the number of underrepresented minorities earning doctorates in the sciences and engineering. Students accepted into the program have exceptional retention rates (95%) and GPAs (3.4) and are broadly distributed in scientific fields. Currently, there are over 200 undergraduates in it. The Center for Women and Information Technology, established at UMBC in 1998, seeks to address and rectify women's under-representation in IT.

The *National Center for the Study of Elections (NCSE)* is located at UMBC in partnership with the Maryland State Board of Elections (SBE). It makes use of the intellectual resources of the University to address issues concerning elections and election and administration in Maryland and across the nation. It provides technical assistance and research support to the SBE in a variety of areas. NCSE developed and maintains the Maryland Voter Information Clearinghouse, a web site for Maryland voters and candidates, and has conducted research on elections technology `www.umbc.edu/mipar/ncse`.

[6] Why can't some penetrator destroy ElectionTracker's entire database and then recreate it in order but with some inconvenent data omitted? Wouldn't that yield a self-consistent new set of timestamps? Yes, but if we require posters to digitally sign the timestamps the signed timestamps, will not be synthesizable in this way. And ElectionTracker could give posters digitally-signed receipts by return email. Then any such erasure and reconstitution attempt would be instantly detectable by anybody with a signed receipt.

E Letter of cost-sharing support from UMBC CSEE Dept. Chair

UMBC

AN HONORS UNIVERSITY IN MARYLAND

Department of Computer Science and Electrical Engineering
University of Maryland, Baltimore County
1000 Hilltop Circle
Baltimore, MD 21250

Phone: 410-455-3500
Fax: 410-455-3969
http://www.csee.umbc.edu

MEMORANDUM

Date: June 1, 2007

To: Dr. Alan T. Sherman

From: Charles Nicholas, Professor and Chair

Subject: Support for Pew Foundation grant

The CSEE Department is prepared to support your proposal to establish a repository for election results, as described in your grant application to the Pew Charitable Trusts. If this proposal is funded, we will provide office space and computer access for Dr. Warren Smith, and appoint him as a non-tenure-track faculty member at an appropriate rank.

In addition, the department will reduce the teaching load for either you (i.e. Dr. Sherman) or Dr. Kalpakis by one course, without cost to this project, to make more time available to carry out the work.