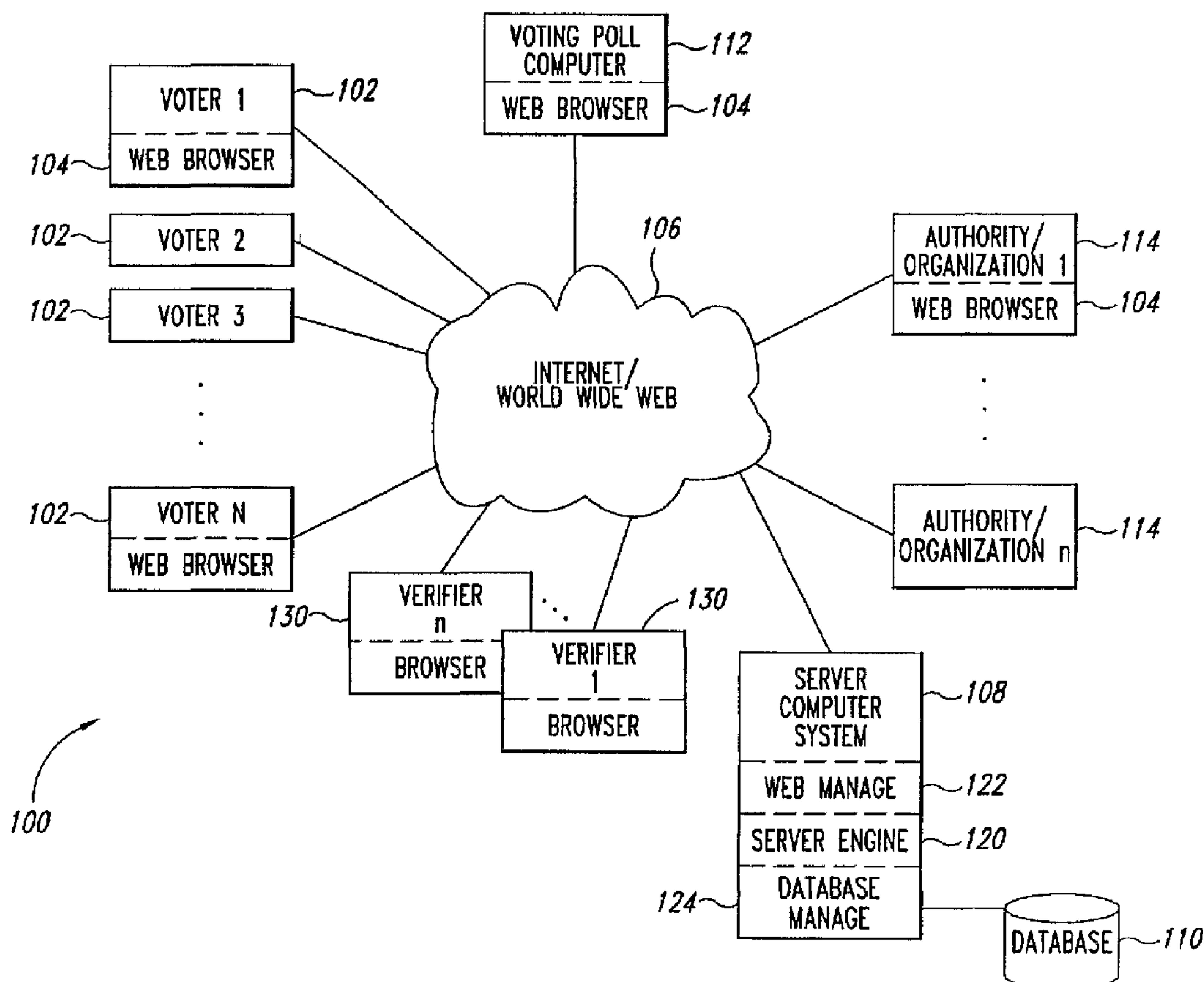




(86) Date de dépôt PCT/PCT Filing Date: 2002/03/25
(87) Date publication PCT/PCT Publication Date: 2002/10/03
(45) Date de délivrance/Issue Date: 2005/05/31
(85) Entrée phase nationale/National Entry: 2003/09/12
(86) N° demande PCT/PCT Application No.: US 2002/009264
(87) N° publication PCT/PCT Publication No.: 2002/077929
(30) Priorités/Priorities: 2001/03/24 (09/816,869) US;
2001/08/09 (60/311,680) US; 2001/08/15 (60/312,671) US;
2001/08/16 (60/313,003) US

(51) Cl.Int.⁷/Int.Cl.⁷ H04L 9/28, G06F 17/60, G07C 13/00,
H04L 12/16, H04L 9/32
(72) Inventeur/Inventor:
NEFF, ANDREW C., US
(73) Propriétaire/Owner:
VOTEHERE, INC., US
(74) Agent: OYEN WIGGS GREEN & MUTALA LLP

(54) Titre : MELANGES SECRETS VERIFIABLES ET LEUR APPLICATION AU VOTE ELECTRONIQUE
(54) Title: VERIFIABLE SECRET SHUFFLES AND THEIR APPLICATION TO ELECTRONIC VOTING



(57) Abrégé/Abstract:

We present a mathematical construct which provides a cryptographic protocol to (verifiably shuffle) a sequence of (k) modular integers, and discuss its application to secure, universally verifiable, multi-authority election schemes. The output of the shuffle



(57) **Abrégé(suite)/Abstract(continued):**

operation is another sequence of (k) modular integers, each of which is the same secret power of a corresponding input element, but the order of elements in the output is kept secret. Though it is a trivial matter for the "shuffler" (who chooses the permutation of the elements to be applied) to compute the output from the input, the construction is important because it provides a linear size proof of correctness for the output sequence (i.e. a proof that it is of the form claimed) that can be checked by one or more arbitrary verifiers. The protocol is shown to be honest-verifier zeroknowledge in a special case, and is computational zeroknowledge in general. On the way to the final result, we also construct a generalization of the well known Chaum-Pedersen protocol for knowledge of discrete logarithm equality ([3], [2]). In fact, the generalization specializes (exactly) to the Chaum-Pedersen protocol in the case $(k)=2$. This result may be of interest on its own. An application to electronic voting is given that matches the features of the best current protocols with significant efficiency improvements. An alternative application to electronic voting is also given that introduces an entirely new paradigm for achieving (Universally Verifiable) elections.

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
3 October 2002 (03.10.2002)

PCT

(10) International Publication Number
WO 02/077929 A2

- (51) International Patent Classification⁷: **G07C 13/00**, G06F 17/60
- (74) Agents: **DALEY-WATSON, Christopher, J.** et al.; Perkins Coie LLP, 1201 Third Avenue, Seattle, WA 98101 (US).
- (21) International Application Number: PCT/US02/09264
- (81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.
- (22) International Filing Date: 25 March 2002 (25.03.2002)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
- | | | |
|------------|-----------------------------|----|
| 09/816,869 | 24 March 2001 (24.03.2001) | US |
| 60/311,680 | 9 August 2001 (09.08.2001) | US |
| 60/312,671 | 15 August 2001 (15.08.2001) | US |
| 60/313,003 | 16 August 2001 (16.08.2001) | US |
- (84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).
- (71) Applicant (*for all designated States except US*): **VOTE-HERE, INC.** [US/US]; Suite 250, 3101 Northup Way, Bellevue, WA 98004 (US).
- (72) Inventor; and
- (75) Inventor/Applicant (*for US only*): **NEFF, Andrew, C.** [US/US]; Suite 250, 3101 Northup Way, Bellevue, WA 98004 (US).
- Published:**
— *without international search report and to be republished upon receipt of that report*
- For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*

(54) Title: VERIFIABLE SECRET SHUFFLES AND THEIR APPLICATION TO ELECTRONIC VOTING

(57) Abstract: We present a mathematical construct which provides a cryptographic protocol to (*verifiably shuffle*) a sequence of (*k*) modular integers, and discuss its application to secure, universally verifiable, multi-authority election schemes. The output of the shuffle operation is another sequence of (*k*) modular integers, each of which is the same secret power of a corresponding input element, but the order of elements in the output is kept secret. Though it is a trivial matter for the "shuffler" (who chooses the permutation of the elements to be applied) to compute the output from the input, the construction is important because it provides a linear size proof of correctness for the output sequence (i.e. a proof that it is of the form claimed) that can be checked by one or more arbitrary verifiers. The protocol is shown to be honest-verifier zeroknowledge in a special case, and is computational zeroknowledge in general. On the way to the final result, we also construct a generalization of the well known Chaum-Pedersen protocol for knowledge of discrete logarithm equality ([3], [2]). In fact, the generalization specializes (*exactly*) to the Chaum-Pedersen protocol in the case (*k*)=2. This result may be of interest on its own. An application to electronic voting is given that matches the features of the best current protocols with significant efficiency improvements. An alternative application to electronic voting is also given that introduces an entirely new paradigm for achieving (*Universally Verifiable*) elections.



WO 02/077929 A2

VERIFIABLE SECRET SHUFFLES AND THEIR APPLICATION TO ELECTRONIC VOTING

5 Field of Invention

The invention relates to electronic voting and particularly to the application of verifiable secret shuffles to electronic voting.

10 Background

The notion of a shuffle of a collection of objects, records, or tokens is simple and intuitive, and useful examples abound in various daily human activities. A gambler in a casino knows that among the cards in his hand, each will be one of 52 unique values, and that
15 no one else at the table will have duplicates of the ones he holds. He does not, however, have any knowledge of how the cards are distributed, even though he may have recorded the exact card order before they were shuffled by the dealer.

In the context of electronic data, the problem of achieving the same kind of random, yet verifiable permutation of an input sequence is surprisingly difficult. The problem is that
20 the data itself is either always visible to the auditor, or it isn't. If it is, then the correspondence between input records and output records is trivial to reconstruct by the auditor, or other observer. If it isn't, then input and output records must be different representations of the same underlying data. But if the output is different enough (that is, encrypted well enough) that the auditor cannot reconstruct the correspondence, then how can
25 the auditor be sure that the shuffler did not change the underlying data in the process of shuffling?

Summary of the Invention

30 One embodiment of the invention provides an electronic voting system for use with a computerized network. The electronic voting system comprises at least one server computer

coupled to receive requests from at least first, second and third voter computers coupled to the computerized network.

The first voter computer is configured to receive, from the server computer, a series of electronic credentials corresponding to an aggregation of electronic credentials received from a plurality of voter computers. Each electronic credential is a Digital Signature Algorithm (DSA) or ElGamal pair. The first voter computer is configured to apply a secret, one-way cryptographic transformation using at least a first secret key to anonymously shuffle the series of electronic credentials, and produce a first shuffled series of credentials for the server computer. Only the first voter computer knows a correspondence between the first series of shuffled credentials and the series of electronic credentials. The first voter computer is further configured to provide a first linear size proof of correctness for the first series of shuffled credentials based on an iterated logarithmic multiplication proof, and to provide a first ballot with a first associated credential.

The second voter computer is configured to receive, from the server computer, the first series of shuffled credentials, to apply the cryptographic transformation using at least a second secret key to anonymously shuffle the first series of shuffled credentials, and produce a second series of shuffled credentials for the server computer. Only the second voter computer knows a correspondence between the first series of shuffled credentials and the second series of shuffled credentials. The second voter computer is further configured to provide to the server computer a second linear size proof of correctness for the second series of shuffled credentials based on the iterated logarithmic multiplication proof with a second ballot having a second associated credential.

The third voter computer is configured to receive the second series of shuffled credentials, to apply the cryptographic transformation using at least a third secret key to anonymously shuffle the second series of shuffled credentials, and produce a third series of shuffled credentials for the server computer. Only the third voter computer knows a correspondence between the second series of shuffled credentials and the third series of shuffled credentials. The third voter computer is further configured to provide a third linear size proof of correctness for the third series of shuffled credentials based on the iterated logarithmic multiplication proof with a third ballot having a third associated credential.

The server computer is configured to receive the proofs of correctness from the first, second and third voter computers and to verify a correctness of the shuffled credentials, and to receive, verify and tally the first, second and third ballots having the respective first, second and third associated credentials.

5 Further aspects of the invention and features of specific embodiments are described below.

Brief Description of the Drawings

10 In drawings which illustrate non-limiting embodiments of the invention:

Figure 1 shows an electronic voting system according to one embodiment of the invention; and,

Figure 2 is a schematic diagram illustrating a method according to one embodiment of the invention.

15

Description

Figure 1 and the following discussion provide a brief, general description of a suitable computing environment in which aspects of the invention can be implemented. Although not
20 required, embodiments of the invention may be implemented as computer-executable instructions, such as routines executed by a general-purpose computer, such as a personal computer or web server. Those skilled in the relevant art will appreciate that aspects of the invention (such as small elections) can be practiced with other computer system configurations, including Internet appliances, hand-held devices, wearable computers,
25 personal digital assistants ("PDAs"), multiprocessor systems, microprocessor-based or programmable consumer electronics, network PCs, mini computers, cell or mobile phones, set-top boxes, mainframe computers, and the like. Aspects of the invention can be embodied in a special purpose computer or data processor that is specifically programmed, configured or constructed to perform one or more of the computer-executable instructions explained
30 herein. Indeed, the term "computer," as generally used herein, refers to any of the above devices, as well as any data processor.

The invention can also be practiced in distributed computing environments where tasks or modules are performed by remote processing devices, which are linked through a communications network, such as a Local Area Network (LAN), Wide Area Network (WAN), or the Internet. In a distributed computing environment, program modules or sub-routines may be located in both local and remote memory storage devices. The invention described herein may be stored or distributed on computer-readable media, including magnetic and optically readable and removable computer disks, stored as firmware in chips, as well as distributed electronically over the Internet or other networks (including wireless networks). Those skilled in the relevant art will recognize that portions of the protocols described herein may reside on a server computer, while corresponding portions reside on client computers. Data structures and transmission of data particular to such protocols are also encompassed within the scope of the invention.

Unless described otherwise, the construction and operation of the various blocks shown in Figure 1 are of conventional design. As a result, such blocks need not be described in further detail herein, as they will be readily understood by those skilled in the relevant art. Referring to Figure 1, a suitable environment of system 100 includes one or more voter or client computers 102, each of which includes a browser program module 104 that permits the computer to access and exchange data with the Internet, including web sites within the World Wide Web portion 106 of the Internet. The voter computers 102 may include one or more central processing units or other logic processing circuitry, memory, input devices (*e.g.*, keyboards, microphones, touch screens, and pointing devices), output devices (*e.g.*, display devices, audio speakers and printers), and storage devices (*e.g.*, fixed, floppy, and optical disk drives), all well known but not shown in Figure 1. The voter computers 102 may also include other program modules, such as an operating system, one or more application programs (*e.g.*, word processing or spread sheet applications), and the like. As shown in Figure 1, there are N number of voter computers 102, representing voters 1, 2, 3 . . . N.

A server computer system 108 or "vote collection center," coupled to the Internet or World Wide Web ("Web") 106, performs much or all of the ballot collection, storing and other processes. A database 110, coupled to the server computer 108, stores much of the web pages and data (including ballots and shuffle validity proofs) exchanged between the voter computers 102, one or more voting poll computers 112 and the server computer 108. The

voting poll computer 112 is a personal computer, server computer, mini-computer, or the like, positioned at a public voting location to permit members of the public, or voters who may not have ready access to computers coupled to the Internet 106, to electronically vote under the system described herein. Thus, the voter computers 102 may be positioned at individual
 5 voter's homes, where one or more voting poll computers 112 are located publicly or otherwise accessible to voters in a public election. The voting poll computer 112 may include a local area network (LAN) having one server computer and several client computers or voter terminals coupled thereto via the LAN to thereby permit several voters to vote simultaneously or in parallel. Note also that the term "voter" is generally used herein to refer to any
 10 individual or organization that employs some or all of the protocols described herein.

Under an alternative embodiment, the system 100 may be used in the context of a private election, such as the election of corporate officers or board members. Under this embodiment, the voter computers 102 may be laptops or desktop computers of shareholders, and the voting poll computer 112 can be one or more computers positioned within the
 15 company (e.g., in the lobby) performing the election. Thus, shareholders may visit the company to access the voting poll computer 112 to cast their votes.

One or more authority or organization computers 114 are also coupled to the server computer system 108 via the Internet 106. If a threshold cryptosystem is employed, then the authority computers 114 each hold a key share necessary to decrypt the electronic ballots
 20 stored in the database 110. Threshold cryptographic systems require that a subset t of the total number of authorities n (*i.e.*, $t < n$) agree to decrypt the ballots, to thereby avoid the requirement that all authorities are needed for ballot decryption. In other words, the objective of a threshold cryptosystem is to share a private key, s , among n members of a group such that messages can be decrypted when a substantial subset, T , cooperate - a (t, n) threshold
 25 cryptosystem. Protocols are defined to (1) generate keys jointly among the group, and (2) decrypt messages without reconstructing the private key. The authority computers 114 may provide their decryption share to the server computer system 108 after the voting period ends so that the server computer system may decrypt the ballots and tally the results.

Furthermore, under the depicted embodiment, each of the authority computers may
 30 perform one shuffle of the ballots, as described herein. In conjunction with each shuffle, each authority computer generates a shuffle validity proof, which may be encrypted and forwarded

to the server computer 108, or stored locally by the authority computer. In an alternative embodiment, an additional set of authority computers are provided, where one set of authority computers shuffle the encrypted ballots and generate shuffle validity proofs, while the second set of authority computers employ keys shares for decrypting the ballots.

5 One or more optional verifier computers 130 may also be provided, similar to the authority computers 114. The verifier computers may receive election transcripts to verify that the election has not been compromised. For example, the verifier computers may receive the shuffle validity proofs from each of the authority computers, as described herein. The verifier computers may perform verifications after the election, and need not be connected to
10 the Internet. Indeed, the verifications may be performed by other computers shown or described herein.

The server, verifier or authority computers may perform voter registration protocols, or separate registration computers may be provided (not shown). The registration computers may include biometric readers for reading biometric data of registrants, such as fingerprint
15 data, voice fingerprint data, digital picture comparison, and other techniques known by those skilled in the relevant art. Voter registration and issuing anonymous certificates for use with verifiable shuffles is described below.

The server computer 108 includes a server engine 120, a web page management component 122, a database management component 124, as well as other components not
20 shown. The server engine 120 performs, in addition to standard functionality, portions of an electronic voting protocol. The encryption protocol may be stored on the server computer, and portions of such protocol also stored on the client computers, together with appropriate constants. Indeed, the above protocol may be stored and distributed on computer readable media, including magnetic and optically readable and removable computer disks, microcode
25 stored on semiconductor chips (*e.g.*, EEPROM), as well as distributed electronically over the Internet or other networks. Those skilled in the relevant art will recognize that portions of the protocol reside on the server computer, while corresponding portions reside on the client computer. Data structures and transmission of data particular to the above protocol are also encompassed within the present invention. Thus, the server engine 120 may perform all
30 necessary ballot transmission to authorized voters, ballot collection, verifying ballots (*e.g.*, checking digital signatures and passing verification of included proofs of validity in ballots),

vote aggregation, ballot decryption and/or vote tabulation. Under an alternative embodiment, the server engine 120 simply collects all electronic ballots as a data collection center. The electronic ballots are then stored and provided to a third party organization conducting the election, such as a municipality, together with tools to shuffle ballots, decrypt the tally and
5 produce election results. Likewise, election audit information, such as shuffle validity proofs and the like may be stored locally or provided to a municipality or other organization.

The web page component 122 handles creation and display or routing of web pages such as an electronic ballot box web page, as described below. Voters and users may access the server computer 108 by means of a URL associated therewith, such as

10 <http://www.votehere.net>, or a URL associated with the election, such as a URL for a municipality. The municipality may host or operate the server computer system 108 directly, or automatically forward such received electronic ballots to a third party vote authorizer who may operate the server computer system. The URL, or any link or address noted herein, can be any resource locator.

15 The web page management process 122 and server computer 108 may have secure sections or pages that may only be accessed by authorized people, such as authorized voters or system administrators. The server computer 108 may employ, e.g., a secure socket layer ("SSL") and tokens or cookies to authenticate such users. Indeed, for small elections, or those where the probability of fraud is low (or results of fraud are relatively inconsequential), the
20 system 100 may employ such simple network security measures for gathering and storing votes as explained below, rather than employing complex electronic encrypted ballots. Methods of authenticating users (such as through the use of passwords), establishing secure transmission connections, and providing secure servers and web pages are known to those skilled in the relevant art.

25 The election scheme and system may use a "bulletin board" where each posting is digitally signed and nothing can be erased. The bulletin board is implemented as a web server. The "ballot box" resides on the bulletin board and holds all of the encrypted ballots. Erasing can be prevented by writing the web server data to a write-once, read-many (WORM) permanent storage medium or similar device.

30 Note that while one embodiment of the invention is described herein as employing the Internet to connect computers, other alternative embodiments are possible. For example,

aspects of the invention may be employed by stand alone computers. Aspects of the invention may also be employed by any interconnected data processing machines. Rather than employing a browser, such machines may employ client software for implementing aspects of the methods or protocols described herein.

5 Referring to Figure 2, a schematic diagram illustrates a basic application of the shuffle protocol to an election, shown as a method 200. In block 202, three encrypted ballots are submitted, one each for voters Joe Smith, Sally Jones, and Ian Kelleigh. In block 204, the list or roll of voters is separated from the encrypted ballots, which are shown in block 206. Thereafter, a one-way reencryption of the ballots is performed to produce a shuffled set of
10 ballots, shown in block 208. A shuffle validity proof is generated based on this first shuffle, shown in block 210. The shuffle validity proof allows a third party to ensure that all input data (the ballots) had the same operation applied to them, and that no altering of the ballots had been performed.

A second shuffle of the (previously shuffled) ballots is performed, to generate a
15 second shuffled set of ballots, shown as block 212. Again, a shuffle validity proof is generated, shown in block 214. The shuffled ballots of block 212 are shuffled a third time, to produce a final shuffled set of ballots under block 216. A third validity proof 218 is likewise generated based on the third shuffle. In sum, a three-by-three shuffle array is provided under this example. Following, the shuffling, the ballots are decrypted to produce a tally, shown as
20 block 220. A third party may verify that the election by analyzing, among other things, each shuffle validity proof to ensure that each shuffler has preserved election integrity.

The shuffle protocol is presented above as effectively separate subroutines that may be employed for various applications, such as in a electronic voting scheme. A first subroutine provides the functionality of scaled, iterated, logarithmic multiplication proofs between a
25 prover and a verifier. A second subroutine provides the functionality of a simple shuffle protocol and employs the scaled, iterated, logarithmic multiplication proofs. Thereafter, a third subroutine implements general shuffle functionality, where the shuffler does not know the exponents, building upon the second subroutine of the simple shuffle. A fourth subroutine extends the third subroutine to shuffling k tuples of elements. Other routines are
30 of course also provided.

One skilled in the art will appreciate that the concepts of the invention can be used in various environments other than the Internet. For example, the concepts can be used in an electronic mail environment in which electronic mail ballots, transactions, or forms are processed and stored. In general, a web page or display description (*e.g.*, the bulletin board) may be in HTML, XML or WAP format, email format, or any other format suitable for displaying information (including character/code based formats, bitmapped formats and vector based formats). Also, various communication channels, such as local area networks, wide area networks, or point-to-point dial-up connections, may be used instead of the Internet. The various transactions may also be conducted within a single computer environment, rather than in a client/server environment. Each voter or client computer may comprise any combination of hardware or software that interacts with the server computer or system. These client systems may include television-based systems, Internet appliances, mobile phones/PDA's and various other consumer products through which transactions can be performed.

In general, as used herein, a "link" refers to any resource locator identifying a resource on the network, such as a display description of a voting authority having a site or node on the network. In general, while hardware platforms, such as voter computers, terminals and servers, are described herein, aspects of the invention are equally applicable to nodes on the network having corresponding resource locators to identify such nodes.

Unless the context clearly requires otherwise, throughout the description and the claims, the words 'comprise,' 'comprising,' and the like are to be construed in an inclusive sense as opposed to an exclusive or exhaustive sense; that is to say, in the sense of "including, but not limited to." Words using the singular or plural number also include the plural or singular number, respectively. Additionally, the words "herein," "above," "below" and words of similar import, when used in this application, shall refer to this application as a whole and not to any particular portions of this application.

The above description of illustrated embodiments of the invention is not intended to be exhaustive or to limit the invention to the precise form disclosed. While specific embodiments of, and examples for, the invention are described herein for illustrative purposes, various equivalent modifications are possible within the scope of the invention, as those skilled in the relevant art will recognize. The teachings of the invention provided

herein can be applied to other encryption applications, not only the electronic voting system described above. For example, the protocol has applications in electronic commerce where both anonymity and auditability are requirements. Examples of this are electronic payment schemes ("e-cash").

5 These and other changes can be made to the invention in light of the above detailed description. In general, in the claims which follow this description, the terms used should not be construed to limit the invention to the specific embodiments disclosed in the specification and the claims, but should be construed to include all encryption systems and methods that operate under the claims to provide data security. Accordingly, the invention is not limited
10 by the disclosure, but instead the scope of the invention is to be determined entirely by the claims.

 While certain aspects of the invention are presented below in certain claim forms, the inventor contemplates the various aspects of the invention in any number of claim forms. For example, while only one aspect of the invention is recited as embodied in a
15 computer-readable medium, other aspects may likewise be embodied in computer-readable medium. Accordingly, the inventor reserves the right to add additional claims after filing the application to pursue such additional claim forms for other aspects of the invention.

 The following description is intended to explain some of the mathematics behind certain aspects of the invention:

The following description is devoted to giving an efficient (linear) method for solving the problem of achieving a random yet verifiable permutation of an input sequence in an important context – ElGamal, or Diffie-Hellman encrypted data. In order to make the exposition as clear and concise as possible, the majority of the following description explicitly refers to the specific case where the operations are carried out in a prime subgroup of $\mathbf{Z}_p^*$, the multiplicative group of units modulo a large prime, p . However, the only property of the underlying (multiplicative) group that is used is that the associated Diffie-Hellman problem is intractable. Thus, the shuffle protocol is also useful when the ElGamal cryptosystem is implemented over other groups such as elliptic curves.

The general Boolean proof techniques of R. Cramer, I. Damgrd, B. Schoenmakers, “Proofs of partial knowledge and simplified design of witness hiding protocols,” *Advances in Cryptology - CRYPTO '94*, Lecture Notes in Computer Science, pp. 174-187, Springer-Verlag, Berlin, 1994, (hereinafter “Cramer et al. (1994)”) and A. De Santis, G. Di Crescenzo, G. Persiano and M. Yung, “On Monotone Formula Closure of SZK,” *FOCS 94*, pp. 454-465, (hereinafter “Santis et al.”) can also be used to construct a proof with the same properties, however, the resulting proof size (complexity) is quadratic, or worse, in the size of the input sequence.

The technique described herein also offers several advantages over the cut-and-choose technique used in K. Sako, J. Kilian, “Receipt-free mix-type voting scheme – A practical solution to the implementation of a voting booth,” *Advances in Cryptology - EUROCRYPT '95*, Lecture Notes in Computer Science, Springer-Verlag, 1995. (hereinafter “Sako et al.”). In this approach, the size of proof is dependent on the probability of a cheating prover that is required to satisfy all participants. In the shuffle protocol described herein, this cheating probability is essentially k/q , where k is the number of elements to be shuffled, and q is the size of the subgroup of $\mathbf{Z}_p^*$ in which the elements are encrypted. Although no analysis of the proof size dependence on cheating probability is done in Sako et al., it appears that, in order to obtain similarly low cheating probability, it will need to be orders of magnitude larger than the size of the proof given herein. (Moreover, if the protocol of Sako et al. is implemented non-interactively, the cheating probability would need to be chosen exceedingly small, because a malicious participant might use considerable off-line computation to generate a forged proof

by exhaustive search. This of course, could be the case with the protocol described herein as well, but the probability k/q is, for all practical values of k and q , certainly small enough – even for offline attacks.)

The results described herein provide for several ways to implement a *universally verifiable* election protocol. In this context, it is worth comparing the elegant homomorphic election protocol of R. Cramer, R. Gennaro, B. Schoenmakers, “A secure and optimally efficient multi-authority election scheme,” *Advances in Cryptology - EUROCRYPT '97*, Lecture Notes in Computer Science, Springer-Verlag, 1997, (hereinafter “Cramer et al. (1997)”). That protocol works well when ballots have only questions of a simple “choose (at most) m of n ” type. This effectively precludes “write-in” responses, as well as “proportional type” questions where the voter is expected to indicate answers in preferential order, and questions are tabulated in accordance with this preference. (Theoretically, proportional type questions can be handled by mapping each valid permutation of selections to a single yes/no response. However, in practice this is infeasible unless the number of choices is quite small.) A couple of somewhat less important disadvantages of the scheme of Cramer et al. (1997) are that it expands vote data size considerably, and that it requires a voter *validity proof*. This proof further expands the vote data size by about an order of magnitude, and is unattractive from a practical perspective, because it presumes special purpose code to be running on the voter’s computer.

The shuffle protocols described herein are constructed entirely from elementary arithmetic operations. They are thus simple to implement, and are imminently practical for the anonymous credential application described herein.

1 Applications to voting

The voting application that occurs immediately is that which employs the usual tabulation/mixing center approach to provide anonymity. In this setting, the protocols described herein offer important advantages. They are much more efficient, and allow the mixing centers to be completely independent of the authorities who hold some share of the key necessary to decrypt ballots.

Perhaps, however, a more valuable and exciting application of the new protocol is for

creating “anonymous credentials”. A member of an authorized group, identified only by a set of DSA, or Diffie-Hellman public keys, can authenticate group membership, and/or sign in a one time way, without revealing his/her *individual* identity. This leads to a novel solution to the voting problem that is universally verifiable, but does not require any special set of “authorities” in order to tabulate. It also offers a better privacy model to the voter, and speeds tabulation enormously since ballots do not need to be encrypted/decrypted. In effect, instead of mixing encrypted vote cyphertexts *after* ballots have been received at the vote collection center, voter credentials are mixed *before* the start of the election. This mixing can naturally be done by the voters themselves to achieve “anonymous authentication”. (See section 7.1.) (It should be noted that the mixing could also be done by a set of authorities, thus providing a more efficient means to implement a threshold privacy election. One where, again, ballots do not need to be encrypted/decrypted.)

2 Notation

In the following, unless explicitly stated otherwise, n will be a positive integer, p and q will be prime integers, publicly known. Arithmetic operations are performed in the modular ring $\mathbf{Z}_p$ (or occasionally $\mathbf{Z}_n$), and $g \in \mathbf{Z}_p$ will have (prime) multiplicative order q . (So, trivially, $q \mid (p-1)$.) In each proof protocol, P will be the prover (shuffler) and V the verifier (auditor).

Recall the Chaum-Pedersen proof of equality for discrete logarithms: For $G, X, H, Y \in \mathbf{Z}_p$ this is a proof of knowledge for the relation

$$\log_G X = \log_H Y \quad (1)$$

It is not known to be zero-knowledge, however it is known to be honest-verifier zero-knowledge. In the next section, a natural multi-variable generalization of this protocol, which also has these properties, will be described. These are sufficient for an application where the verifier is implemented via the Fiat-Shamir heuristic. (See A. Fiat, A. Shamir, “How to prove yourself: Practical solutions to identification and signature problems,” *Advances in Cryptology - CRYPTO '86*, Lecture Notes in Computer Science, pp. 186-194, Springer-Verlag, New York, 1987, (hereinafter “Fiat et al.”) and Cramer et al. (1997).)

Definition 1 *An instance of this proof, as above, will be denoted by*

$$\mathcal{CP}(G, X, H, Y).$$

Definition 2 *For fixed $g \in \mathbf{Z}_p^*$, let $\otimes_g$ be the binary operator on $\langle g \rangle \times \langle g \rangle$ defined by*

$$\log_g(x \otimes_g y) = \log_g x \log_g y$$

for all $x, y \in \langle g \rangle$. Alternatively

$$g^a \otimes_g g^b = g^{ab} = (g^a)^b = (g^b)^a$$

for all $a, b \in \mathbf{Z}_q$. Following the conventions used for summations and multiplications, the notation

$$\bigotimes_{i=1}^k X_i = X_0 \otimes_g X_1 \otimes_g \cdots \otimes_g X_k$$

is also used. This operation is referred to as logarithmic multiplication base, g .

In each of the notations in the preceding definition, the subscript g may be omitted when its value is clear from context.

Remark 1 Notice that

$$\log_G X = \log_H Y \iff G \otimes_g Y = H \otimes_g X \quad (2)$$

The following collection of well known results will be used in the following description.

Lemma 1 *Let $f(x) \in \mathbf{Z}_q[x]$, be a polynomial of degree d . Then there are at most d values $z_1, \dots, z_d \in \mathbf{Z}_q$ such that $f(z_i) = 0$.*

Corollary 1 *Let $f(x), g(x) \in \mathbf{Z}_q[x]$ be two monic polynomials of degree at most d , with $f \neq g$. Then there are at most $d - 1$ values $z_1, \dots, z_{d-1} \in \mathbf{Z}_q$ such that $f(z_i) = g(z_i)$.*

Corollary 2 Let $f(x), g(x) \in \mathbf{Z}_q[x]$ be two monic polynomials of degree at most d , with $f \neq g$. If $t \in_R \mathbf{Z}_q$ (t is selected at random from $\mathbf{Z}_q$), then

$$P(\{t : f(t) = g(t)\}) \leq \frac{d-1}{q}$$

Corollary 3 Let $f(x), g(x) \in \mathbf{Z}_q[x]$ be any two polynomials of degree at most d . Then for every constant $R \neq 0$, there are at most d values, $z_1(R), \dots, z_d(R)$, such that $f(z_i(R)) = Rg(z_i(R))$.

Definition 3 Let $f(x)$ be a polynomial in $\mathbf{Z}_q[x]$. We denote by χ_f the (unordered) set of all roots of f .

$$\chi_f \doteq \{t \in \mathbf{Z}_q : f(t) = 0\} \quad (3)$$

Definition 4 If $\Lambda \subset \mathbf{Z}_q$, and $R \in \mathbf{Z}_q$, we write

$$R\Lambda \doteq \{t \in \mathbf{Z}_q : t = Ru, u \in \Lambda\} \quad (4)$$

Corollary 4 Let $f(x), g(x) \in \mathbf{Z}_q[x]$ be any two polynomials of degree at most d . Fix constants, $R \neq 0$, $\gamma \neq 0$, and $\delta \neq 0$. If $t \in_R \mathbf{Z}_q$, then

$$P(\{t : f(\gamma t) = Rg(\delta t)\}) \leq \frac{d}{q}$$

Lemma 2 Let $\mathbf{Z}_q^k$ be the standard k -dimensional vector space over $\mathbf{Z}_q$, and fix $v = (v_1, \dots, v_k) \in \mathbf{Z}_q^k$, $v \neq 0$, and $a \in \mathbf{Z}_q$. If $r \in_R \mathbf{Z}_q^k$ is chosen at random, then

$$P(\{r : v \cdot r = a\}) = \frac{1}{q}$$

3 Proofs for iterated logarithmic multiplication

For the rest of this section, all logarithmic multiplications will be computed relative to a fixed element g , and hence the subscript will be omitted in notation. The following problem is fundamental to the shuffle protocols which are to come later.

Iterated Logarithmic Multiplication Problem: Two sequences $\{X_i\}_{i=1}^k$ and $\{Y_i\}_{i=1}^k$ are publicly known. The prover, $\mathcal{P}$, also knows $u_i = \log_g X_i$ and $v_i = \log_g Y_i$ for all i , but these are unknown to the verifier, $\mathcal{V}$. $\mathcal{P}$ is required to convince $\mathcal{V}$ of the relation

$$\bigotimes_{i=1}^k X_i = \bigotimes_{i=1}^k Y_i \quad (5)$$

without revealing any information about the secret logarithms u_i and v_i .

The protocol described herein is precisely a higher dimensional generalization of the Chaum-Pedersen protocol discussed above. In fact, in the case $k = 2$, the protocol is *exactly* the Chaum-Pedersen protocol. The presentation will be considerably simplified by restricting the problem instance to a case where

$$X_i \neq 1, Y_i \neq 1 \quad \forall 1 \leq i \leq k \quad (6)$$

Clearly, if any of these inequalities do not hold, then there is no sense in constructing a proof since equation (5) can be seen to hold or not by inspection. (If $X_i = 1$ then $x_i = 0$ and so equation (5) holds if and only if $Y_j = 1$ for some j . Similarly with the roles of X and Y reversed.)

Iterated Logarithmic Multiplication Proof Protocol (ILMPP) :

1. $\mathcal{P}$ secretly generates, randomly and independently from $\mathbf{Z}_q$, $k - 1$ elements, $\theta_1, \dots, \theta_{k-1}$. $\mathcal{P}$ then computes

$$\begin{aligned} A_1 &= Y_1^{\theta_1} \\ A_2 &= X_2^{\theta_1} Y_2^{\theta_2} \\ &\vdots \\ A_i &= X_i^{\theta_{i-1}} Y_i^{\theta_i} \\ &\vdots \\ A_k &= X_k^{\theta_{k-1}} \end{aligned} \quad (7)$$

and reveals to $\mathcal{V}$ the sequence $A_1, \dots, A_k$.

2. $\mathcal{V}$ generates a random challenge, $\gamma \in \mathbf{Z}_q$ and reveals it to $\mathcal{P}$.
3. $\mathcal{P}$ computes $k - 1$ elements, $r_1, \dots, r_{k-1}$, of $\mathbf{Z}_q$ satisfying

$$\begin{aligned}
 Y_1^{r_1} &= A_1 X_1^{-\gamma} \\
 X_2^{r_1} Y_2^{r_2} &= A_2 \\
 \vdots &= \vdots \\
 X_i^{r_{i-1}} Y_i^{r_i} &= A_i \\
 \vdots &= \vdots \\
 X_k^{r_{k-1}} &= A_k Y_k^{(-1)^{(k-1)}\gamma}
 \end{aligned} \tag{8}$$

and reveals the sequence $r_1, \dots, r_{k-1}$ to $\mathcal{V}$. (It will be shown in the proof of completeness, below, how these values are computed.)

4. $\mathcal{V}$ accepts the proof if and only if all of the equations in (8) hold.

Theorem 1 *The ILMPP is a three-move, public coin proof of knowledge for the relationship in equation (5) which is special honest-verifier zero-knowledge. The number of exponentiations required to construct the proof is k , and the number of exponentiations required to verify it is $2k$. If $\mathcal{V}$ generates challenges randomly, the probability of a forged proof is $1/q$.*

Remark 2 Note that in constructing the proof, all exponentiations can be done to the same base, g , so fixed base algorithms can be employed. (See A.J. Menezes, P.C. van Oorschot, and S.A. Vanstone. Handbook of Applied Cryptography, CRC Press, 1997, p. 623.)

Proof: The protocol is clearly three-move and public coin. The exponentiation count in the construction of the proof looks like it should be $2k - 2$, but actually it can be constructed with only k exponentiations. This is because $\mathcal{P}$ knows the logarithms x_i and y_i , and hence can compute A_i as $A_i = g^{\theta_{i-1}x_i + \theta_i y_i}$ for all $2 \leq i \leq k - 1$.

Completeness

Completeness means that, given arbitrary $\vec{\theta} = (\theta_1, \dots, \theta_{k-1})$ and γ , $\mathcal{P}$ can always find $\vec{r} = (r_1, \dots, r_{k-1})$ satisfying the system of equations in (8). To see that this is the case, take $\log_g$ of each side of the equations in (8), and set $\bar{r}_i = r_i - \theta_i$ for $1 \leq i \leq k-1$. One obtains the following $k \times (k-1)$ system of linear equations in $\mathbf{Z}_q$ for $\bar{r}_1, \dots, \bar{r}_{k-1}$

$$\begin{pmatrix} y_1 & 0 & 0 & 0 & \cdots & 0 \\ x_2 & y_2 & 0 & 0 & \cdots & 0 \\ 0 & x_3 & y_3 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 0 & x_{k-1} & y_{k-1} \\ 0 & 0 & \cdots & 0 & 0 & x_k \end{pmatrix} \begin{pmatrix} \bar{r}_1 \\ \bar{r}_2 \\ \bar{r}_3 \\ \vdots \\ \bar{r}_{k-2} \\ \bar{r}_{k-1} \end{pmatrix} = \begin{pmatrix} -\gamma x_1 \\ 0 \\ 0 \\ \vdots \\ 0 \\ (-1)^{(k-1)} \gamma y_k \end{pmatrix} \quad (9)$$

The $(k-1) \times (k-1)$ sub-system

$$\begin{pmatrix} x_2 & y_2 & 0 & 0 & \cdots & 0 \\ 0 & x_3 & y_3 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 0 & x_{k-1} & y_{k-1} \\ 0 & 0 & \cdots & 0 & 0 & x_k \end{pmatrix} \begin{pmatrix} \bar{r}_1 \\ \bar{r}_2 \\ \vdots \\ \bar{r}_{k-2} \\ \bar{r}_{k-1} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ \gamma y_k \end{pmatrix} \quad (10)$$

is non-singular since its determinant is $\prod_{i=2}^k x_i$, which is non-zero by assumption (6). Hence, one can always solve it for $\bar{r}_1, \dots, \bar{r}_{k-1}$. In fact, the solution is

$$\bar{r}_i = (-1)^{(k-i-1)} \gamma \prod_{j=i+1}^k \left(\frac{y_j}{x_j} \right) \quad (11)$$

However, under the hypotheses of the problem, (10) actually implies (9). This is because

$$\begin{aligned} 0 &= \prod_{i=1}^k x_i - \prod_{i=1}^k y_i \\ &= \begin{vmatrix} x_1 & y_1 & 0 & 0 & 0 & \cdots & 0 \\ 0 & x_2 & y_2 & 0 & 0 & \cdots & 0 \\ 0 & 0 & x_3 & y_3 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & x_{k-1} & y_{k-1} \\ (-1)^k y_k & 0 & 0 & \cdots & 0 & 0 & x_k \end{vmatrix} \end{aligned} \quad (12)$$

which, combined with the fact that the sub-matrix on the left of equation (10) is non-singular, means that the first column vector of the $k \times k$ matrix in (12) must be a linear combination of the remaining $k - 1$ column vectors.

Soundness

If the first column vector of the matrix on the left of equation (12) is not a linear combination of the remaining $k - 1$ column vectors, then there can be *at most* one value of $\gamma \in \mathbb{Z}_q$ for which equation (9) holds. Thus, if γ is chosen randomly, there is at most a chance of 1 in q that $\mathcal{P}$ can produce $r_1, \dots, r_{k-1}$ which convince $\mathcal{V}$.

Special Honest-Verifier Zero-knowledge

Honest-verifier zero-knowledge holds because, for random γ and random $\vec{r} = (r_1, \dots, r_{k-1})$, and for

$$\vec{A} = \begin{pmatrix} A_1 \\ A_2 \\ \vdots \\ A_{k-1} \\ A_k \end{pmatrix} = \begin{pmatrix} X_1^\gamma Y_1^{r_1} \\ X_2^{r_1} Y_2^{r_2} \\ \vdots \\ X_{k-1}^{r_{k-2}} Y_{k-1}^{r_{k-2}} \\ X_k^{r_{k-1}} Y_k^{(-1)^k \gamma} \end{pmatrix} \quad (13)$$

the triple $(\vec{A}, \gamma, \vec{r})$ is an accepting conversation. Hence, a simulation is constructed by generating random and independent γ and r_i , and setting $\vec{A}$ as in equation (13).

It is easy to see that the distribution so generated for $\vec{A}$ is identical to that generated according to (7), again because the first column vector of the matrix in (12) is a fixed linear combination of the remaining column vectors. So if $\mathcal{V}$ is honest, the simulation is perfect. Since the challenge, γ , can be chosen freely, there is also special honest-verifier zero-knowledge.

Remark 3 The solutions for $\bar{r}_i$ in (11) could also be written formally as

$$\bar{r}_i = (-1)^{(i-1)} \gamma \prod_{j=1}^i \left(\frac{x_j}{y_j} \right)$$

However, this will not work if some of the y_j are 0. In the case of equation (11), this problem was avoided by assumption (6). Of course, the main part of the solution could just have well been set up under the assumption that $y_i \neq 0$ for all $1 \leq i \leq k - 1$ – the choice of expression for r_i just needs to be kept consistent with the form of the assumption.

Remark 4 In the case $k = 2$, the ILMPP reduces exactly to the well known Chaum-Pedersen protocol.

Remark 5 Special Soundness

As is the case with the Chaum-Pedersen protocol, which proves that $\mathcal{P}$ knows $s = \log_G X = \log_H Y$, the ILMPP proves that $\mathcal{P}$ knows $s_1, \dots, s_k$ such that $\log_g X_i = \log_g Y_i$, and $\prod_{i=1}^k s_i = 1$. This is clear because from two accepting conversations, $(\vec{A}, \gamma, \vec{r})$ and $(\vec{A}, \gamma', \vec{r}')$, with the same first move and $\gamma \neq \gamma'$, a witness,

$$(w, \rho) = (\gamma - \gamma', \vec{r} - \vec{r}')$$

can be extracted satisfying

$$\begin{aligned} Y_1^{\rho_1} &= X_1^{-w} \\ Y_2^{\rho_2} &= X_2^{-\rho_1} \\ \vdots &= \vdots \\ Y_i^{\rho_i} &= X_i^{-\rho_{i-1}} \\ \vdots &= \vdots \\ Y_k^{(-1)^k w} &= X_k^{-\rho_{k-1}} \end{aligned} \tag{14}$$

Since $w = \gamma - \gamma' \neq 0$, it follows that $\rho_i \neq 0$ for all $1 \leq i \leq k-1$. Thus, with an appropriate small modification to the statement of the problem, it satisfies special soundness.

4 The Simple k -Shuffle

The first shuffle proof protocol constructed herein requires a restrictive set of conditions. It will be useful for two reasons. First, it is a basic building block of the more general shuffle proof protocol to come later. Fortuitously, it also serves a second important purpose. A single instance of this proof can be constructed to essentially “commit” a particular permutation. This can be important when shuffles need to be performed on tuples of $\mathbf{Z}_p$ elements, which is exactly what is required when shuffling ElGamal pairs, as in the voting application.

Definition 5 Suppose two sequences of k elements of $\mathbf{Z}_p$, $X_1, \dots, X_k$, and $Y_1, \dots, Y_k$, along with two additional elements, C and D , are publicly known. Suppose also, that the prover, $\mathcal{P}$, knows $x_i = \log_g X_i$, $y_i = \log_g Y_i$, $c = \log_g C$, and $d = \log_g D$, but that all these values are unknown to the verifier, $\mathcal{V}$. $\mathcal{P}$ is required to convince $\mathcal{V}$ that there is some permutation, $\pi \in \Sigma_k$, with the property that

$$Y_i^d = X_{\pi(i)}^c \quad (15)$$

for all $1 \leq i \leq k$ without revealing any information about x_i , y_i , π , c , or d .

This problem is referred to herein as the *Simple k -Shuffle Problem*.

Remark 6 For this section, and the remainder of this description, the following simplifying assumptions have been made, that in all shuffle constructions

1. $x_i \neq x_j$ for $i \neq j$ (and hence, of course, $y_i \neq y_j$ for $i \neq j$).
2. $x_i \neq 1$ for all $1 \leq i \leq k$.

There are obvious ways to handle these special cases. Moreover, in practice, they will “essentially never” occur since elements are usually random.

The protocol of the previous section, in combination with corollary 2, provide the tools necessary to solve this problem in a fairly straightforward manner.

Simple k -Shuffle Proof Protocol:

1. $\mathcal{V}$ generates a random $t \in \mathbf{Z}_q$ and gives it to $\mathcal{P}$ as a challenge.
2. $\mathcal{P}$ and $\mathcal{V}$ publicly compute $U = D^t = g^{dt}$, $W = C^t = g^{ct}$,

$$\vec{X} = (\hat{X}_1, \dots, \hat{X}_k) = (X_1/U, \dots, X_k/U)$$

and

$$\vec{Y} = (\hat{Y}_1, \dots, \hat{Y}_k) = (Y_1/W, \dots, Y_k/W)$$

3. $\mathcal{P}$ and $\mathcal{V}$ execute the ILMPP for the two length $2k$ vectors

$$\begin{aligned}\Phi &= (\vec{X}, \overbrace{C, C, \dots, C}^k) \\ \Psi &= (\vec{Y}, \overbrace{D, D, \dots, D}^k)\end{aligned}\tag{16}$$

The protocol succeeds ($\mathcal{V}$ accepts the proof) if and only if $\mathcal{V}$ accepts this ILMPP.

Theorem 2 *The Simple k -Shuffle Proof Protocol is a four-move, public coin proof of knowledge for the relationship in equation (15). It satisfies special soundness, and is special honest-verifier zero-knowledge. The number of exponentiations required to construct the proof is $2k$, and the number of exponentiations required to verify it is $4k$.*

If $\mathcal{V}$ generates challenges randomly, the probability of a forged proof is less than or equal to

$$(k-1)/q + (q-k+1)/q^2 = (qk - q + q - k + 1)/q^2 < k/q$$

Remark 7 The observations of remark 2 also apply in this case.

Proof: All of the required properties follow immediately from the results of the previous section. (Special soundness can be argued from remark 5.) A forged proof can only be generated in two conditions.

1. The challenge t is one of the special values for which

$$\prod_{i=1}^k (t - x_i/d) = \prod_{i=1}^k (t - y_i/c)$$

2. The challenge t is not one of the special values in 1 above, *and* the ILMPP is forged.

By corollary 2, the probability of 1 is at most $(k-1)/q$, and the probability of 2 is $(q-k+1)/q^2$ by the results of the previous section.

4.1 A complexity improvement

Both the size and complexity of the simple k -shuffle protocol can be improved by a factor of 2. Instead of corollary 2, corollary 4 is used. One may be inclined to replace the k copies of D and k copies of C in equation (16) with single entries g^{d^k} and g^{c^k} respectively. Unfortunately, this would ruin the zero-knowledge property of the protocol. Instead, the protocol is modified as follows.

Simple k -Shuffle Proof Protocol II:

1. $\mathcal{P}$ generates randomly and independently β from $\mathbf{Z}_q$ and τ from $\mathbf{Z}_q - \{0\}$, computes

$$\begin{aligned} B &= g^\beta \\ T &= g^\tau \end{aligned} \tag{17}$$

and reveals B and T to $\mathcal{V}$.

2. $\mathcal{V}$ generates a random λ from $\mathbf{Z}_q$ and reveals it to $\mathcal{P}$.
3. $\mathcal{P}$ computes s by

$$s = \beta + \lambda\tau - \prod_{i=1}^k \left(\frac{x_i}{y_i} \right) = \beta + \lambda\tau - \left(\frac{d}{c} \right)^k \tag{18}$$

and reveals s to $\mathcal{V}$.

4. $\mathcal{V}$ generates a random $t \in \mathbf{Z}_q$ and gives it to $\mathcal{P}$ as a challenge.
5. $\mathcal{P}$ and $\mathcal{V}$ publicly compute $U = D^t = g^{dt}$, $W = C^t = g^{ct}$,

$$\vec{\hat{X}} = (\hat{X}_1, \dots, \hat{X}_k) = (X_1/U, \dots, X_k/U)$$

and

$$\vec{\hat{Y}} = (\hat{Y}_1, \dots, \hat{Y}_k) = (Y_1/W, \dots, Y_k/W)$$

6. $\mathcal{P}$ secretly generates, randomly and independently from $\mathbf{Z}_q$, k elements, $\theta_1, \dots, \theta_k$. $\mathcal{P}$ then computes

$$\begin{aligned}
 A_1 &= \hat{Y}_1^{\theta_1} \\
 A_2 &= \hat{X}_2^{\theta_1} \hat{Y}_2^{\theta_2} \\
 &\vdots \\
 A_i &= \hat{X}_i^{\theta_{i-1}} \hat{Y}_i^{\theta_i} \\
 &\vdots \\
 A_k &= \hat{X}_k^{\theta_{k-1}} \hat{Y}_k^{\theta_k} \\
 A_{k+1} &= g^{\theta_k}
 \end{aligned} \tag{19}$$

and reveals to $\mathcal{V}$ the sequence $A_1, \dots, A_{k+1}$.

7. $\mathcal{V}$ generates a random challenge, $\gamma \in \mathbf{Z}_q$ and reveals it to $\mathcal{P}$.
8. $\mathcal{P}$ computes k elements, $r_1, \dots, r_k$, of $\mathbf{Z}_q$ satisfying

$$\begin{aligned}
 \hat{Y}_1^{r_1} &= A_1 \hat{X}_1^{-\gamma} \\
 \hat{X}_2^{r_1} \hat{Y}_2^{r_2} &= A_2 \\
 &\vdots \\
 \hat{X}_i^{r_{i-1}} \hat{Y}_i^{r_i} &= A_i \\
 &\vdots \\
 \hat{X}_k^{r_{k-1}} \hat{Y}_k^{r_k} &= A_k \\
 g^{r_k} &= A_{k+1} \left(BT^\lambda g^{-s} \right)^{(-1)^k \gamma}
 \end{aligned} \tag{20}$$

and reveals the sequence $r_1, \dots, r_k$ to $\mathcal{V}$.

9. $\mathcal{V}$ accepts the proof if and only if all of the equations in (20) hold.

Theorem 3 *Simple k -Shuffle Proof Protocol II is a five-move, public coin proof of knowledge for the relationship in equation (15). It satisfies special soundness, and is special honest-verifier zero-knowledge. The number of exponentiations required to construct the proof is $k+4$, and the number of exponentiations required to verify it is $2k+2$.*

If $\mathcal{V}$ generates challenges randomly, the probability of a forged proof remains less than or equal to

$$(k-1)/q + (q-k+1)/q^2 = (qk - q + q - k + 1)/q^2 < k/q$$

Proof Sketch: All of the arguments are very similar, property by property, to the arguments constructed in the case of the original protocol. The main difference is that one makes an appeal to corollary 4 rather than corollary 2.

Definition 6 An instance of the Simple k -Shuffle Protocol II is denoted herein by $\mathcal{SS}_k(\vec{X}, \vec{Y}, C, D)$, where $\vec{X} = (X_1, \dots, X_k)$, $\vec{Y} = (Y_1, \dots, Y_k)$, C , and D are as in definition 5.

5 The General k -Shuffle

An obvious limitation of the simple k -Shuffle protocol is that the shuffler, $\mathcal{P}$, must know all the original exponents $x_1, \dots, x_k$ and $y_1, \dots, y_k$. In many applications, this will not be the case. The goal of this section is to eliminate that restriction.

General k -Shuffle Problem: Two sequences of k elements of $\mathbf{Z}_p$, $X_1, \dots, X_k$, and $Y_1, \dots, Y_k$ are publicly known. In addition, constants $c, d \in \mathbf{Z}_q$ are known only to $\mathcal{P}$, but commitments $C = g^c$ and $D = g^d$ are made public. $\mathcal{P}$ is required to convince $\mathcal{V}$ that there is some permutation, $\pi \in \Sigma_k$, with the property that

$$Y_i^d = X_{\pi(i)}^c \tag{21}$$

for all $1 \leq i \leq k$ without revealing any information about π , c , or d .

General k -Shuffle Proof Protocol:

(For the sake of presentation, consider the case $d = 1$. The general case can be reduced to this case simply by changing the group generator from g to C , and making the substitution $c = c/d$.)

1. For $1 \leq i \leq k$, $\mathcal{P}$ generates a_i , b_i , u_i and w_i randomly and independently from $\mathbf{Z}_q$, and computes

$$\begin{aligned} A_i &= g^{a_i} \\ B_i &= g^{b_i} \\ U_i &= g^{u_i} \\ W_i &= g^{w_i} \end{aligned} \tag{22}$$

In addition, $\mathcal{P}$ generates γ from $\mathbf{Z}_q - \{0\}$, and also z_0 , z_1 randomly and independently from $\mathbf{Z}_q$, and computes

$$\begin{aligned} \Gamma &= g^\gamma \\ Z_0 &= g^{z_0} \\ Z_1 &= g^{z_1} \end{aligned} \tag{23}$$

and, for $1 \leq i \leq k$

$$C_{\pi(i)} = A_i^\gamma = g^{\gamma a_i} \tag{24}$$

and finally

$$\begin{aligned} \Lambda &= \frac{\prod_{i=1}^k Y_i^{w_i}}{K_1^c K_0^{\gamma c} \prod_{i=1}^k X_i^{\gamma c u_i}} \\ &= g^{-(cz_1 + \gamma cz_1)} \prod_{i=1}^k X_i^{\gamma c u_i - cw_{\pi^{-1}(i)}} \end{aligned} \tag{25}$$

$\mathcal{P}$ then reveals the ordered sequences A_i , B_i , C_i , U_i , and W_i along with X_0 , Y_0 , and Λ to $\mathcal{V}$.

2. For $1 \leq i \leq k$, $\mathcal{V}$ chooses e_i randomly and independently from $\mathbf{Z}_q - \{0\}$ and returns the sequence to $\mathcal{P}$.
3. $\mathcal{P}$ computes

$$D_i = g^{\gamma e_{\pi(i)} b_{\pi(i)}} \tag{26}$$

for $1 \leq i \leq k$ and reveals these to $\mathcal{V}$.

4. $\mathcal{V}$ generates $t \in \mathbb{Z}_q - \{0\}$ and returns it to $\mathcal{P}$ as a challenge.
5. For $1 \leq i \leq k$, $\mathcal{P}$ secretly computes the exponents

$$\begin{aligned} r_i &= a_i + te_i b_i \\ s_i &= \gamma r_{\pi(i)} \end{aligned} \tag{27}$$

6. $\mathcal{P}$ and $\mathcal{V}$ then execute the simple k -shuffle, $\mathcal{SS}_k(\vec{R}, \vec{S}, g, \Gamma)$, where

$$\begin{aligned} \vec{R} &= (R_1, \dots, R_k) = (g^{r_1}, \dots, g^{r_k}) \\ \vec{S} &= (S_1, \dots, S_k) = (g^{s_1}, \dots, g^{s_k}) \end{aligned} \tag{28}$$

(Note that $\mathcal{P}$ need not explicitly compute R_i and S_i in order to construct the proof, while $\mathcal{V}$ can compute R_i and S_i as $R_i = A_i B_i^{te_i}$ and $S_i = C_i D_i^t$. Thus, this proof protocol requires $\mathcal{P}$ to compute $k+4$ exponentiations, and $\mathcal{V}$ to compute $4k+2$ exponentiations.)

7. $\mathcal{P}$ reveals the exponents

$$\begin{aligned} \rho_i &= u_i + r_i - a_i \\ \sigma_i &= w_i + s_i - c_i \end{aligned} \tag{29}$$

8. $\mathcal{V}$ checks that

$$\begin{aligned} g^{\rho_i} &= U_i R_i / A_i \\ g^{\sigma_i} &= W_i S_i / C_i \end{aligned} \tag{30}$$

and evaluates

$$\begin{aligned} G &= Z_0 \prod_{i=1}^k X_i^{\rho_i} \\ H &= \prod_{i=1}^k Y_i^{\sigma_i} \end{aligned} \tag{31}$$

9. $\mathcal{P}$ computes

$$\begin{aligned} G_0 &= G^\gamma \\ G_1 &= (Z_1 G_0)^c \end{aligned} \tag{32}$$

and executes with $\mathcal{V}$ the two Chaum-Pedersen proofs, $\mathcal{CP}(g, \Gamma, G, G_0)$ and $\mathcal{CP}(g, C, Z_1 G_0, G_1)$.
(Thus proving to $\mathcal{V}$ that equations 32 hold.)

10. $\mathcal{V}$ finally checks that

$$\frac{H}{G_1} = \Lambda \quad (33)$$

11. $\mathcal{V}$ accepts the proof if and only if

- (a) All equations that must be checked (steps 8 and 10) are satisfied
- (b) $\mathcal{V}$ accepts the simple shuffle proof of step 6
- (c) $\mathcal{V}$ accepts both Chaum-Pedersen proofs in step 9

6 The Multi-Authority Voting Application

Much of the setting for the conventional voting application can be found in Cramer et al. (1997). Votes are submitted as ElGamal pairs of the form $(g^{\alpha_i}, h^{\alpha_i} m)$ (or a sequence of these pairs if more data is required), where m is some standard encoding of the voter choices, the α_i are generated secretly by the voters, and h is a public parameter constructed via a dealerless secret sharing scheme [T. Pedersen, “A threshold cryptosystem without a trusted party,” *Advances in Cryptology - EUROCRYPT '91*, Lecture Notes in Computer Science, pp. 522-526, Springer-Verlag, 1991.] (hereinafter “Pedersen”). Once the polls are closed (voting finished), an independent collection of authorities sequentially shuffles the ballots. Each shuffle is constructed as

$$(\bar{X}_i, \bar{Y}_i) = (g^{\beta_{\pi(i)}} X_{\pi(i)}, h^{\beta_{\pi(i)}} Y_{\pi(i)}) \quad (34)$$

where the r_i are randomly chosen from $\mathbf{Z}_q$ and

$$(X_i, Y_i) = (g^{\alpha_i}, h^{\alpha_i} M_i)$$

On output of the final shuffle, the final collection of encrypted ballots is decrypted in accordance with the threshold scheme, and the clear text votes are tabulated in full view by normal election rules.

The authorities who participate in the sequential shuffles, may be arbitrary in number, and they may be completely different from those who hold shares of the election private key. The sequence of ballots which are finally decrypted can only be matched with the original sequence of submitted ballots if *all* of the shuffling authorities collude, since each of their permutations is completely arbitrary.

Although the form of each shuffle is different from the shuffle discussed in section 5, a proof protocol can be constructed with minor modifications. The details are presented in the following subsection.

6.1 Shuffles of ElGamal Pairs

1. For $1 \leq i \leq k$, $\mathcal{P}$ generates a_i , b_i , u_i and w_i randomly and independently from $\mathbf{Z}_q$, and computes

$$\begin{aligned} A_i &= g^{a_i} \\ B_i &= g^{b_i} \\ U_i &= g^{u_i} \\ W_i &= g^{w_i} \end{aligned} \tag{35}$$

In addition, $\mathcal{P}$ generates γ from $\mathbf{Z}_q - \{0\}$, and also x_0 , y_0 and t_0 randomly and independently from $\mathbf{Z}_q$, and computes

$$\begin{aligned} \Gamma &= g^\gamma \\ X_0 &= g^{x_0} \\ Y_0 &= g^{y_0} \end{aligned} \tag{36}$$

and, for $1 \leq i \leq k$

$$C_{\pi(i)} = A_i^\gamma = g^{\gamma a_i} \tag{37}$$

and finally

$$\Lambda = \frac{g^{t_0} X_0^\gamma \prod_{i=1}^k (\bar{Y}_i^{w_i} / Y_i^{\gamma u_i})}{h^{t_0} Y_0^\gamma \prod_{i=1}^k (\bar{X}_i^{w_i} / X_i^{\gamma u_i})} \tag{38}$$

$$= g^{\gamma(x_0-y_0)} \left(\frac{g}{h}\right)^{t_0-\sum_{i=1}^k w_i \beta_{\pi(i)}} \prod_{i=1}^k \left(\frac{X_i}{Y_i}\right)^{\gamma u_i - w_{\pi^{-1}(i)}}$$

$\mathcal{P}$ then reveals the ordered sequences A_i , B_i , C_i , U_i , and W_i along with X_0 , Y_0 , and Λ to $\mathcal{V}$.

2. For $1 \leq i \leq k$, $\mathcal{V}$ chooses e_i randomly and independently from $\mathbf{Z}_q - \{0\}$ and returns the sequence to $\mathcal{P}$.

3. $\mathcal{P}$ computes

$$D_i = g^{\gamma e_{\pi(i)} b_{\pi(i)}} \quad (39)$$

for $1 \leq i \leq k$ and reveals these to $\mathcal{V}$.

4. $\mathcal{V}$ generates $c \in \mathbf{Z}_q - \{0\}$ and returns it to $\mathcal{P}$ as a challenge.

5. For $1 \leq i \leq k$, $\mathcal{P}$ secretly computes the exponents

$$\begin{aligned} r_i &= a_i + c e_i b_i \\ s_i &= \gamma r_{\pi(i)} \end{aligned} \quad (40)$$

6. $\mathcal{P}$ and $\mathcal{V}$ then execute the simple k -shuffle, $\mathcal{SS}_k(\vec{R}, \vec{S}, g, \Gamma)$, where

$$\begin{aligned} \vec{R} &= (R_1, \dots, R_k) = (g^{r_1}, \dots, g^{r_k}) \\ \vec{S} &= (S_1, \dots, S_k) = (g^{s_1}, \dots, g^{s_k}) \end{aligned} \quad (41)$$

(Note that $\mathcal{P}$ need not explicitly compute R_i and S_i in order to construct the proof, while $\mathcal{V}$ can compute R_i and S_i as $R_i = A_i B_i^{c e_i}$ and $S_i = C_i D_i^c$. Thus, this proof protocol requires $\mathcal{P}$ to compute $k+4$ exponentiations, and $\mathcal{V}$ to compute $4k+2$ exponentiations.)

7. $\mathcal{P}$ reveals the exponents

$$\begin{aligned} \rho_i &= u_i + r_i - a_i \\ \sigma_i &= w_i + s_i - c_i \end{aligned} \quad (42)$$

8. $\mathcal{V}$ checks that

$$\begin{aligned} g^{\rho_i} &= U_i R_i / A_i \\ g^{\sigma_i} &= W_i S_i / C_i \end{aligned} \quad (43)$$

and evaluates

$$\begin{aligned} G &= X_0 \prod_{i=1}^k X_i^{\rho_i} \\ H &= Y_0 \prod_{i=1}^k Y_i^{\rho_i} \\ \bar{G} &= \prod_{i=1}^k \bar{X}_i^{\sigma_i} \\ \bar{H} &= \prod_{i=1}^k \bar{Y}_i^{\sigma_i} \end{aligned} \quad (44)$$

9. $\mathcal{P}$ computes

$$\begin{aligned} G_0 &= G^\gamma \\ H_0 &= H^\gamma \end{aligned} \quad (45)$$

and executes with $\mathcal{V}$ the two Chaum-Pedersen proofs, $\mathcal{CP}(g, \Gamma, G, G_0)$ and $\mathcal{CP}(g, \Gamma, H, H_0)$.
(Thus proving to $\mathcal{V}$ that equations 45 hold.)

10. $\mathcal{P}$ and $\mathcal{V}$ compute

$$\begin{aligned} \Delta_0 &= \bar{G} / G_0 \\ \Delta_1 &= \bar{H} / H_0 \end{aligned} \quad (46)$$

11. $\mathcal{P}$ computes

$$t = t_0 + \sum_{i=1}^k \gamma c_i b_i \beta_i \quad (47)$$

and reveals it to $\mathcal{V}$.

12. $\mathcal{V}$ finally checks that

$$\frac{\Delta_1 / h^t}{\Delta_0 / g^t} = \Lambda \quad (48)$$

13. $\mathcal{V}$ accepts the proof if and only if

- (a) All equations that must be checked (steps 8 and 12) are satisfied
- (b) $\mathcal{V}$ accepts the simple shuffle proof of step 6
- (c) $\mathcal{V}$ accepts both Chaum-Pedersen proofs in step 9

6.2 One Pass Tabulation

In the standard mix-net implementation, it is assumed that two distinct phases must take place:

Shuffling (or mixing) The set of encrypted ballots are mixed. This means that the entire encrypted (and iteratively shuffled) ballot box must be passed sequentially from one authority to another until a sufficient number of mixing stages have been performed.

More precisely, let $\mathcal{A}_1, \dots, \mathcal{A}_n$ be the sequence of shuffling, or mixing entities, usually called *authorities*. Let $\mathcal{B}$ be a sequence of encrypted ballots. Sequentially, each $\mathcal{A}_i$ performs the following operations.

1. $\mathcal{A}_i$ receives

$$\mathcal{B} = ((X_1, Y_1), \dots, (X_k, Y_k))$$

from $\mathcal{A}_{i-1}$, along with all required authentication and validity proofs. (In the case $i = 1$, $\mathcal{A}_1$ receives $\mathcal{B}$ directly from the vote collection center.)

2. $\mathcal{A}_i$ performs all necessary authentication checks and validity (proof) verifications.
3. If any of the checks in the step 2 fail, tabulation is aborted, or possibly restarted. Otherwise, $\mathcal{A}_i$ computes

$$\bar{\mathcal{B}} = ((\bar{X}_1, \bar{Y}_1), \dots, (\bar{X}_k, \bar{Y}_k))$$

according to equation 34, and also the corresponding validity proof, P_i , of section 6.1. (Typically, this is a non-interactive version of the protocol.)

4. Renaming $\bar{\mathcal{B}}$ as $\mathcal{B}$, $\mathcal{A}_i$ sends $\mathcal{B}$ to $\mathcal{A}_{i+1}$ along with P_i and any other required authentication or validity checks. In the case $i = n$, the mixing is terminated, and tabulation proceeds to the decryption phase.

Decryption After shuffling (mixing), the resulting set of ballots are all encrypted to the election public key, usually denoted by h . In order to decrypt these ballots, each authority must now contribute a decryption share. This is accomplished as follows.

1. Let

$$\mathcal{B} = ((X_1, Y_1), \dots, (X_k, Y_k))$$

be the set of encrypted ballots output by the final mixing authority, $\mathcal{A}_n$, and let $S_{\mathcal{D}} = \{\mathcal{D}_1, \dots, \mathcal{D}_t\}$ be the subset of decryption authorities participating in the decryption. (The value t is determined by the threshold used when the election public key, h , was created. See Pedersen.)

2. For each $1 \leq j \leq t$, and in parallel:

- (a) $\mathcal{D}_j$ receives the sequence $X_1, \dots, X_k$ along with proper authentication of its validity.
- (b) If a check of the authentication provided in step 2a fails, $\mathcal{D}_j$ should abort the decryption. Otherwise, for each $1 \leq i \leq k$, $\mathcal{D}_j$ computes

$$Z_{ij} = X_i^{z(j, S_{\mathcal{D}})s_j} \quad (49)$$

where s_j is $\mathcal{D}_j$'s decryption share, and $z(j, S_{\mathcal{D}})$ is a publically computable constant multiplier. (See Pedersen.) $\mathcal{D}_j$ also produces a Chaum-Pedersen proof of validity, C_{ij} , for equation 49.

- (c) $\mathcal{D}$ returns all Z_{ij} and C_{ij} to a tabulation center.

3. Assuming that all C_{ij} received at the tabulation center are valid, the set of clear text ballots are computed as

$$M_i = Y_i / \prod_{j=1}^t Z_{ij} \quad (50)$$

4. The election tally is finally computed by counting the M_i as in any election.

If the set, $S_{\mathcal{D}}$, of decryption authorities (or even just some subset of it) is included in the set of shuffling authorities (this is almost always the case in practice), a disadvantage of this procedure is that it requires each authority that is in both sets to receive two separate communications, each of which may be quite large. The disadvantage is particularly severe when the authorities are connected by a slow communication link.

It is observed that it is possible for some or all of the decryption to take place at the same time as mixing. This is accomplished simply by combining, during the shuffling phase, shuffling step 3 with decryption step 2 for each of those shuffling authorities who are members of $S_{\mathcal{D}}$. Those shuffling authorities who are not members of $S_{\mathcal{D}}$ can follow the steps unmodified.

So, if $\mathcal{A}_j$ is a member of $S_{\mathcal{D}}$, in shuffling step 4, rather than just send $\mathcal{B}$ to $\mathcal{A}_{j+1}$, $\mathcal{A}_j$ sends *both*

$$\mathcal{B} = ((X_1, Y_1), \dots, (X_k, Y_k))$$

and

$$\mathcal{Z} = ((Z_{1j}, C_{1j}), \dots, (Z_{kj}, C_{kj}))$$

where Z_{ij} and C_{ij} are as above. Further, instead of using $\mathcal{B}$ as the input to its shuffle, $\mathcal{A}_{j+1}$ should use

$$\mathcal{B} = ((X_1, Y_1/Z_{1j}), \dots, (X_k, Y_k/Z_{kj}))$$

In this way, if all of the decryption authorities are also shuffle authorities, as is usually the case, then the final output will be a set of decrypted ballots, thus eliminating the need for a separate decryption phase. If only $t_0 < t$ decryption authorities are also shuffle authorities, then the final shuffled ballot box will consist of elements that are encrypted with a public key which is shared among the remaining $t - t_0$ decryption authorities. Thus, in this case there is still a need for a decryption phase, but the amount of communication required is reduced by the corresponding amount.

7 k -Shuffles of DSA Public Keys

The general k -shuffle is ideally suited to verifiably permuting a set of DSA, or Diffie-Hellman public keys. By this it is meant that a new set of DSA public keys is produced, which is computationally unlinkable to the original set, but which *verifiably* represents the same set of private keys. This can be extremely valuable when one wishes to anonymize a set of authenticated keys while still protecting the integrity of the original group of private keys – the election setting is just one such example.

The technique is only sketched here, but the details should then be readily apparent. It is assumed that initially all the public keys are of the form (g, H) , $H = g^s$, where g is some fixed generator and s is the private key. That is, loosely, “all the keys use the same base”. The protocol proceeds as follows:

1. Shuffler, or mixer, is presented with g and the list of keys H_i .
2. Shuffler executes the general k -shuffle with $C = g$, and $Y_i = H'_i$ (the new public keys), implementing the verifier’s random challenges via the Fiat-Shamir heuristic. (That is, a non-interactive version of the proof protocol is executed.)
3. Shuffler “returns” the entire proof transcript.
4. Assuming the transcript verifies, set $g = C$, and $H_i = H'_i$. By changing the common base to C , the private keys all remain the same since

$$H = g^s \iff H' = C^s \tag{51}$$

7.1 Anonymous Voters

In the voting application, it is often said that for election integrity one must know “who voted”, but for privacy, one *must not* know “how voted”. The technique of this section solves the privacy/integrity dilemma in a new way. Instead of knowing “who voted”, one only knows that the person who voted is *a member of a set of authorized voters!* As a consequence, this results in a voting solution that

1. Does not require key sharing to implement a distributed trust tabulation scheme.

2. Guarantees *computational privacy* to the voter, rather than *threshold privacy*, which is a necessary evil of other voting solutions based on distributed trust tabulation. (If a threshold number of authorities agree to collude, all voters' votes can be decrypted.)
3. Does not require encryption or decryption of voted ballots.

Of course, one must look after the problem of "double voting", but the technique of this section is easily modified to take care of that as follows.

AV-1. In step 3, the voter (shuffler) – who knows one of the private keys s_0 in this case – *signs* his voted ballot using a DSA signature scheme with group generator (or base), C , and key pair (s_0, H'_0) . (H'_0 is the "post shuffle" public key which belongs to the voter. The voter knows its place in the new sequence, since he/she executed the shuffle. And, moreover, the properties of the shuffle guarantee that $H'_0 = C^{s_0}$ – that is, s_0 is the private key corresponding to H'_0 .)

AV-2. In step 4, assuming that the shuffle transcript checks, and that the ballot signature checks, the vote center simply removes H'_0 from the list of authorized keys, and starts the process again waiting for the next ballot request. The new list of public keys is now one smaller, and unless the voter (shuffler) knew more than one private key in the first place, he/she now knows *none* of the new private keys, and hence can not vote again.

The resulting election protocol is *Universally Verifiable* if all the shuffle transcripts and signatures are maintained.

Although this election scheme offers superior privacy characteristics (the secrecy of each voter's ballot is protected by the voter, rather than by a distributed set of authorities), there is a potential practical drawback. The amount of computation required both of the voter (that is, the voter's computer) and the vote collection server computer for the purpose of authentication is much larger than in standard election protocols, where the voter must only compute a standard public key signature. However, if this additional computation is expected to cause a problem, then the same anonymization can be performed well in advance of the election, before the ballot is available.

This is accomplished by issuing standard PKI certificates issued in response to a certificate request which is anonymously signed exactly as the ballot is signed in step AV-1 above. Of course these certificate requests would contain no identifying personal information, but they would otherwise be completely standard, including, most importantly, a standard public key for which the voter has the corresponding private key. At vote time, the ballot can be simply signed with this ordinary private key. With this strategy, it is possible to re-use the anonymous certificates. Thus the computational cost of the anonymous signatures can be amortized over many elections.

8 k -Shuffles of Tuples

It should be clear that in section 5, the simple shuffle generated essentially “froze” the permutation that could be proved. This makes it easy to see how to extend the previous section to shuffles of k l -tuples of elements of $\langle g \rangle$, or k l -tuples of ElGamal pairs. Thinking of a sequence of k l -tuples as a $k \times l$ array, a single simple k -shuffle can serve to prove that all columns have been permuted according to the same permutation. Precisely, this means that the same values of A_i , B_i , C_i and D_i in both the General k -Shuffle protocol, and the ElGamal Shuffle protocol are used for all columns. (In the General k -Shuffle, the same value of t can be used for all columns, while in the ElGamal Shuffle, the same value of c can be used for all columns, though this is not essential.)

8.1 DSA key shuffles without common base

The observation of this section also allows a generalization of the DSA key shuffle protocol of section 7. Rather than maintaining the entire set of public keys to the same base, $g \leftrightarrow C$, the keys are maintained as independent pairs (g_i, H_i) . The shuffler can pick an arbitrary subset of key pairs, (G_i, H_i) , shuffle them “as 2-tuples”, and return the result. This makes shuffling more manageable if the original set is large, at the cost of increasing the work per key by about 50%.

CLAIMS

I claim:

1. An electronic voting system for use with a computerized network, comprising:
 - at least one server computer coupled to receive requests from at least first, second and third voter computers coupled to the computerized network;
 - wherein the first voter computer is programmed to receive, from the server computer, a series of electronic credentials corresponding to an aggregation of electronic credentials received from a plurality of voter computers, wherein each electronic credential is a Digital Signature Algorithm (DSA) or ElGamal pair, and wherein the first voter computer is programmed to apply a secret, one-way cryptographic transformation using at least a first secret key to anonymously shuffle the series of electronic credentials, and produce a first shuffled series of credentials for the server computer, wherein only the first voter computer knows a correspondence between the first series of shuffled credentials and the series of electronic credentials, and wherein the first voter computer is further programmed to provide a first linear size proof of correctness for the first series of shuffled credentials based on an iterated logarithmic multiplication proof, and to provide a first ballot with a first associated credential;
 - wherein the second voter computer is programmed to receive, from the server computer, the first series of shuffled credentials, to apply the cryptographic transformation using at least a second secret key to anonymously shuffle the first series of shuffled credentials, and produce a second series of shuffled credentials for the server computer, wherein only the second voter computer knows a correspondence between the first series of shuffled credentials and the second series of shuffled credentials, and wherein the second voter computer is further programmed to provide to the server computer a second linear size proof of correctness for the second series of shuffled credentials based on the iterated logarithmic multiplication proof with a second ballot having a second associated credential;
 - wherein the third voter computer is programmed to receive the second series of shuffled credentials, to apply the cryptographic transformation using at least a third secret key to anonymously shuffle the second series of shuffled credentials, and

produce a third series of shuffled credentials for the server computer, wherein only the third voter computer knows a correspondence between the second series of shuffled credentials and the third series of shuffled credentials, and wherein the third voter computer is further programmed to provide a third linear size proof of correctness for the third series of shuffled credentials based on the iterated logarithmic multiplication proof with a third ballot having a third associated credential; and

wherein the server computer is programmed to receive the proofs of correctness from the first, second and third voter computers and to verify a correctness of the shuffled credentials, and to receive, verify and tally the first, second and third ballots having the respective first, second and third associated credentials.

2. The system of claim 1, further comprising:

a plurality of authority computers coupled to the computerized network, wherein the first, second and third ballots are encrypted under a threshold, discrete log asymmetric encryption process using underlying groups Z_p or elliptic curve;

and wherein each of the plurality of authority computers, in turn:

receives at least the first, second and third encrypted ballots,

secretly shuffles the ballots;

applies a decryption key share to the secretly shuffled ballots to

partially decrypt the secretly shuffled ballots; and

passes the partially decrypted and secretly shuffled ballots to a next authority computer.

3. The system of claim 1 wherein the first, second and third voter computers are programmed to provide to the server computer at least an identification of first, second and third public keys in the aggregation of electronic credentials associated with the first, second and third voter computers, all respectively, and

wherein the server computer is further programmed to remove the first, second and third public keys in response thereto.

4. The system of claim 1 wherein the computerized network includes the World Wide Web, wherein each of the first, second and third voter computers include a web browser program.

5. The system of claim 1 wherein at least one of the first, second and third voter computers is a palm-sized computer, cell phone, wearable computer, interactive television terminal or Internet appliance.

6. A computer system for receiving a sequence of elements, comprising a computer coupled to a computer network, the computer having a memory with computer-readable code embodied therein comprising:

code means for receiving a sequence of electronic data elements representing individual data files,

code means for applying a cryptographic transformation using at least a first secret value to anonymously permute the sequence of electronic data elements and produce a first shuffled sequence of electronic data elements, wherein a server computer knows a correspondence between the first shuffled sequence of electronic data elements and the sequence of electronic data elements, and

code means for generating a first linear size proof of correctness for the first shuffled sequence of electronic data elements based on an iterated logarithmic multiplication proof employing a binary operator, wherein each of the data elements in the sequence of data elements is associated with a common base value, and wherein verifying the proof of correctness employs a three-move, public coin proof of knowledge wherein in response to a verifying request, the computer computes a series of exponents by solving an associated series of linear equations.

7. The system of claim 6 wherein the received sequence of electronic data elements are encrypted using Z_p or elliptic curve groups using a key unknown to the computer, the computer-readable code further comprising:

code means for receiving a series of randomly generated values e_i from a verifier computer;

code means for secretly generating a series of values D_i based on a secret, one-way cryptographic transformation that employs the received series of values e_i and secretly generated values;

code means for permuting the sequence of electronic data elements to produce the first shuffled sequence of elements based on the series of values D_i and a secret value; and

code means for providing the values D_i and a series of proof values based on the cryptographic transformation as a proof of knowledge that the server computer has access to how the cryptographic transformation permuted the sequence of electronic data elements to produce the first shuffled sequence of elements without revealing the cryptographic transformation to the verifier computer.

8. The system of claim 6 wherein the computer-readable code further comprises:

code means for receiving a plurality of public keys from a corresponding plurality of individuals, wherein each of the plurality of individuals have a private key corresponding to one of the plurality of public keys;

code means for receiving a request from one of the plurality of individuals having a one private key;

code means for providing at least a subset of the plurality of public keys to the requesting individual;

code means for receiving a file and a shuffle of the plurality of public keys and a linear size proof of correctness for the shuffled public keys based on a iterated logarithmic multiplication proof and a value corresponding to the one private key, wherein the value is associated with the file and provides proof that the one individual has knowledge of the one private key without revealing the one private key;

code means for checking the proof of correctness;

code means for checking that the value is mathematically related to a one of the public keys and the file; and

code means for reducing the plurality of public keys by the one public key.

9. The system of claim 6 wherein the sequence of electronic elements are public keys, and wherein computer-readable code further comprises code means for checking, in response to a request from an individual, that the individual has a value uniquely and mathematically related to a one of the public keys; and

5 if so, issuing a certificate to the one individual.

10. A computer-implemented method, comprising:

receiving a request from one of a plurality of individuals having a private key corresponding to one of a plurality of public keys;

10 providing at least a shuffled subset of the plurality of public keys to the requesting individual;

receiving a file F , and an authentication consisting of:

(1) a new shuffled set of public keys S ,

(2) a shuffle proof of correctness for the new shuffled set of public
15 keys, and

(3) a signature value x from the requesting individual; and

accepting the authentication if:

(a) verification of the shuffle proof of correctness succeeds, and

(b) $V(F, x, s_0)$ is true, wherein V is a signature verification function
20 and s_0 is one public key from the new shuffled set of public keys S .

11. A computer-implemented method, comprising:

receiving a request from a computer associated with one private key corresponding to one public key in a plurality of public keys, wherein each of the
25 plurality of public keys corresponds to one of a plurality of private keys;

providing at least a shuffled subset of the plurality of public keys to the requesting computer;

receiving a file from the requesting computer;

receiving a new shuffled subset of the plurality of public keys and a linear size
30 proof of correctness for the new shuffled subset of public keys based on an iterated logarithmic multiplication proof and a value corresponding to the one private key,

wherein the value is associated with the file and provides proof that the requesting computer has knowledge of the one private key without revealing the one private key;
 checking the proof of correctness; and
 checking that the value is mathematically related, in a substantially unique
 5 way, to the one public key and the file.

12. The method of claim 11, further comprising receiving from the requesting computer at least an indication of the one public key in the new shuffled subset of public keys.

10 13. The method of claim 11, further comprising removing the one public key from the new shuffled subset of public keys.

14. The method of claim 11, further comprising:

15 receiving from each of a plurality of authorities, in sequence, a shuffled set of the plurality of public keys H' based on a secret cryptographic shuffle operation performed on at least a subset of the plurality of public keys H to produce the shuffled set of the plurality of public keys H';

receiving from each of a plurality of authorities, in sequence, a verification transcript of the cryptographic shuffle operation; and

20 verifying a correctness of the cryptographic shuffle operation based on the verification transcript; and if verified, then setting the shuffled set of the plurality of public keys from H' to H.

15. The method of claim 11, further comprising:

25 at a time after receiving at least some of the plurality of public keys, setting at least a subset of the then received plurality of public keys to a received shuffled set of the plurality of public keys, wherein the shuffled set of the plurality of public keys have been received from a third party.

30 16. The method of claim 11, further comprising:

following the checking steps, issuing a certificate;

thereafter, receiving the issued certificate from one of a plurality of individuals; and

providing an electronic ballot to the one individual.

5 17. The method of claim 11, further comprising digitally signing a received request to produce a public key infrastructure ("PKI") certificate.

18. The method of claim 11, further comprising:

receiving issued certificates from at least some of a plurality of user computers
10 and providing initial electronic ballots in response thereto; and

receiving unencrypted voted ballots from the at least some of the plurality of user computers.

15 19. A computer-implemented cryptographic method between a prover computer and a verifier computer, the method comprising:

secretly shuffling a first series of electronic data elements with respect to a second series of electronic data elements to produce a shuffled set;

generating a linear size proof of correctness for the shuffled set based on proving that a first set of roots for a first polynomial equation is equal to a fixed
20 constant multiple of a second set of roots for a second polynomial equation; and

wherein the linear size proof of correctness that the first set of roots is equal to a fixed multiple of the second sets of roots is obtained by: specifying a second fixed constant, evaluating the first polynomial equation at a first random point to obtain a first value, evaluating the second polynomial equation at the first random point to
25 obtain a second value, and determining that the second value is equal to a product of the first value and the second fixed constant.

20. The method of claim 19 wherein at least some of the elements in the first sequence of elements are electronic ballots.

30

21. The method of claim 19 wherein at least some of the elements in the shuffled set are ElGamal pairs.

22. The method of claim 19, further comprising repeating the secretly shuffling,
5 generating and providing for l-tuple of elements in an input sequence of elements.

23. The method of claim 19 wherein secretly shuffling the first series of elements includes receiving a subset of a set of identifying elements, wherein each identifying element in the set corresponds to an individual, and wherein the method further comprises:

10 receiving an anonymous certificate if the verifying computer verifies the proof of correctness.

24. An electronic voting system for use with a computerized network, comprising:

15 a plurality of authority computers coupled to the computerized network to receive a plurality of ballots encrypted under a threshold, discrete log asymmetric encryption process;

wherein each of the plurality of authority computers comprises a memory with computer-readable code embodied therein comprising code means for, in turn:

receiving at least the plurality of encrypted ballots,

20 secretly shuffling the ballots;

applying a decryption key share to the secretly shuffled ballots to partially decrypt the secretly shuffled ballots;

generating a proof of validity for the shuffled and partially decrypted ballots; and

25 passing the partially decrypted and secretly shuffled ballots to a next authority computer; and

wherein a last of the authority computers comprises code means for applying a last decryption key share to the secretly shuffled ballots to fully decrypt the secretly shuffled ballots.

30

25. The system of claim 24 wherein each of the plurality of authority computers further comprises code means for substantially simultaneously:

(1) secretly shuffling the ballots,

(2) applying the decryption key share to the secretly shuffled ballots to
5 partially decrypt the secretly shuffled ballots, and

(3) generating the proof of validity for the partially decrypted ballots.

26. A computer-readable medium whose contents provide instructions, when implemented by a computer, perform a shuffling of a sequence of electronic data elements,
10 comprising:

providing a request from a computer associated with one private key
corresponding to one public key in a plurality of public keys, wherein each of the
plurality of public keys corresponds to one of a plurality of private keys;

receiving at least a shuffled subset of the plurality of public keys;

15 providing a file;

producing a new shuffled subset of the plurality of public keys and a proof of
correctness for the new shuffled subset of public keys based on an iterated logarithmic
multiplication proof and a value corresponding to the one private key, wherein the
value is associated with the file and provides proof of knowledge of the one private
20 key without revealing the one private key.

27. The computer-readable medium of claim 26 wherein the computer-readable medium
is located in a logical node in a computer network receiving the sequence of electronic data
elements and the contents.

28. The computer-readable medium of claim 26 wherein the computer-readable medium
is a computer-readable disk.

29. The computer-readable medium of claim 26 wherein the computer-readable medium
30 is a memory of a computer system.

30. The computer-readable medium of claim 26 wherein the proof of correctness is a linear size proof of correctness that includes computing first and second polynomial equations, and wherein an evaluation of the first and second polynomials at a randomly chosen point results in a same computed value.

5

31. A carrier wave embodying a computer data signal representing sequences of statements and instructions which, when implemented by a computer, perform a shuffling of a sequence of electronic data elements, comprising:

10 providing a request from a computer associated with one private key corresponding to one public key in a plurality of public keys, wherein each of the plurality of public keys corresponds to one of a plurality of private keys;

receiving at least a shuffled subset of the plurality of public keys;

providing a file;

15 producing a new shuffled subset of the plurality of public keys and a proof of correctness for the new shuffled subset of public keys based on an iterated logarithmic multiplication proof and a value corresponding to the one private key, wherein the value is associated with the file and provides proof of knowledge of the one private key without revealing the one private key.

20 32. The carrier wave of claim 31 wherein the carrier wave is located on an internet connection link to a voting authority server computer.

33. A computer-implemented method for performing a mixing of electronic data elements, comprising:

25 receiving at least an original subset of multiple public keys, wherein each of public keys in the original subset of public keys corresponds to one private key in a set of private keys; and

mixing the original subset of public keys to produce a mixed set of public keys,

wherein the mixed set of public keys can not be correlated with the original subset of public keys, but wherein the mixed set of public keys corresponds to the set of private keys.

5 34. The method of claim 33, further comprising:
 providing a file; and
 producing a proof of correctness for the new mixed subset of public keys
 based on an iterated logarithmic multiplication proof and a value corresponding to the
 one private key, wherein the value is associated with the file and provides proof of
10 knowledge of the one private key without revealing the one private key.

 35. The method of claim 33, further comprising generating a proof of correctness for the
 mixed set of public keys, wherein the proof of correctness proves that the mixed set of public
 keys corresponds to the set of private keys.

15 36. The method of claim 33 wherein the mixed subset of public keys correspond to a
 sequence of ballots.

 37. The method of claim 33, further comprising generating a proof of correctness for the
20 mixed set of public keys that includes computing first and second polynomial equations,
 wherein an evaluation of the first and second polynomials at a randomly chosen point results
 in a same computed value.

25

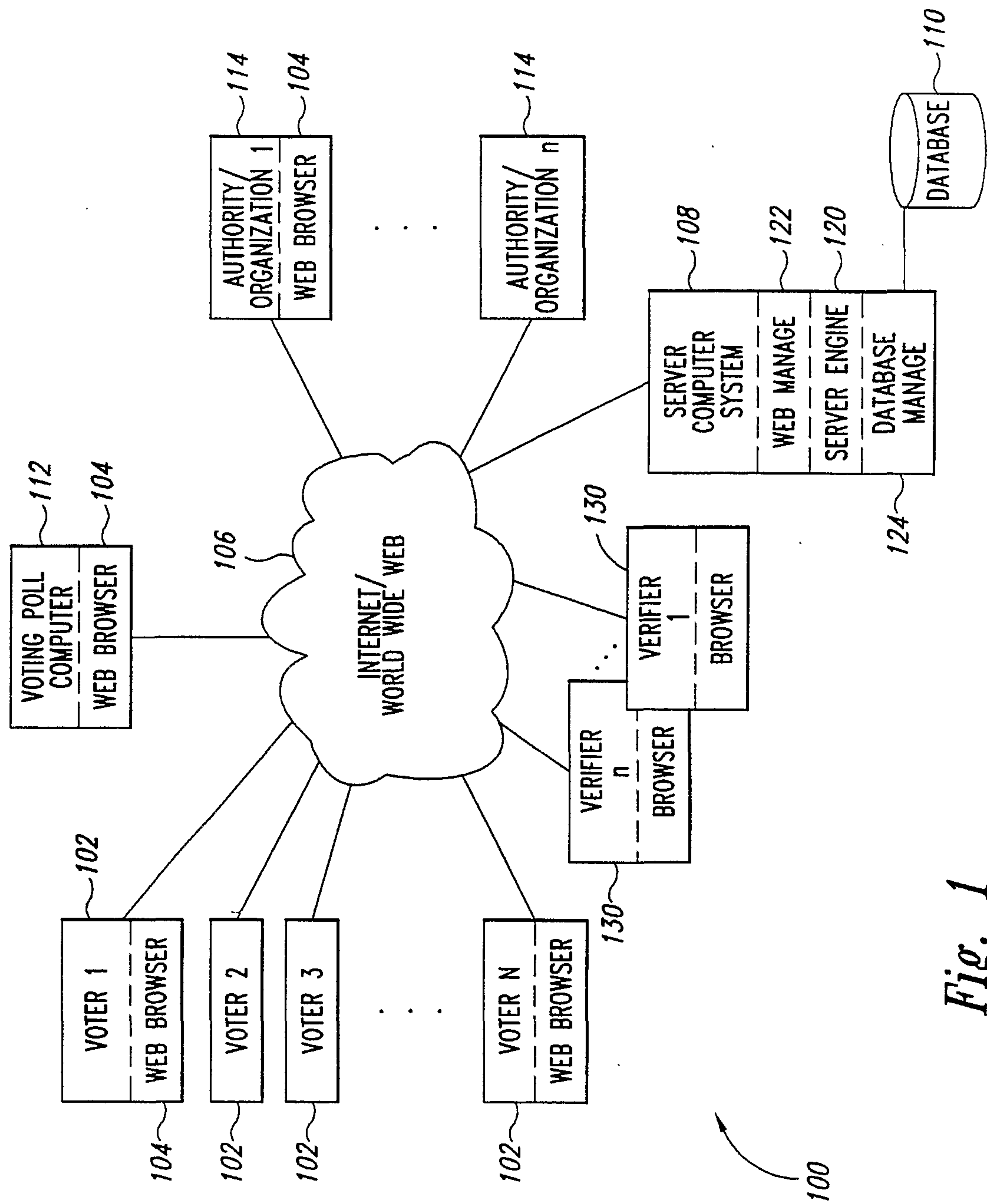


Fig. 1

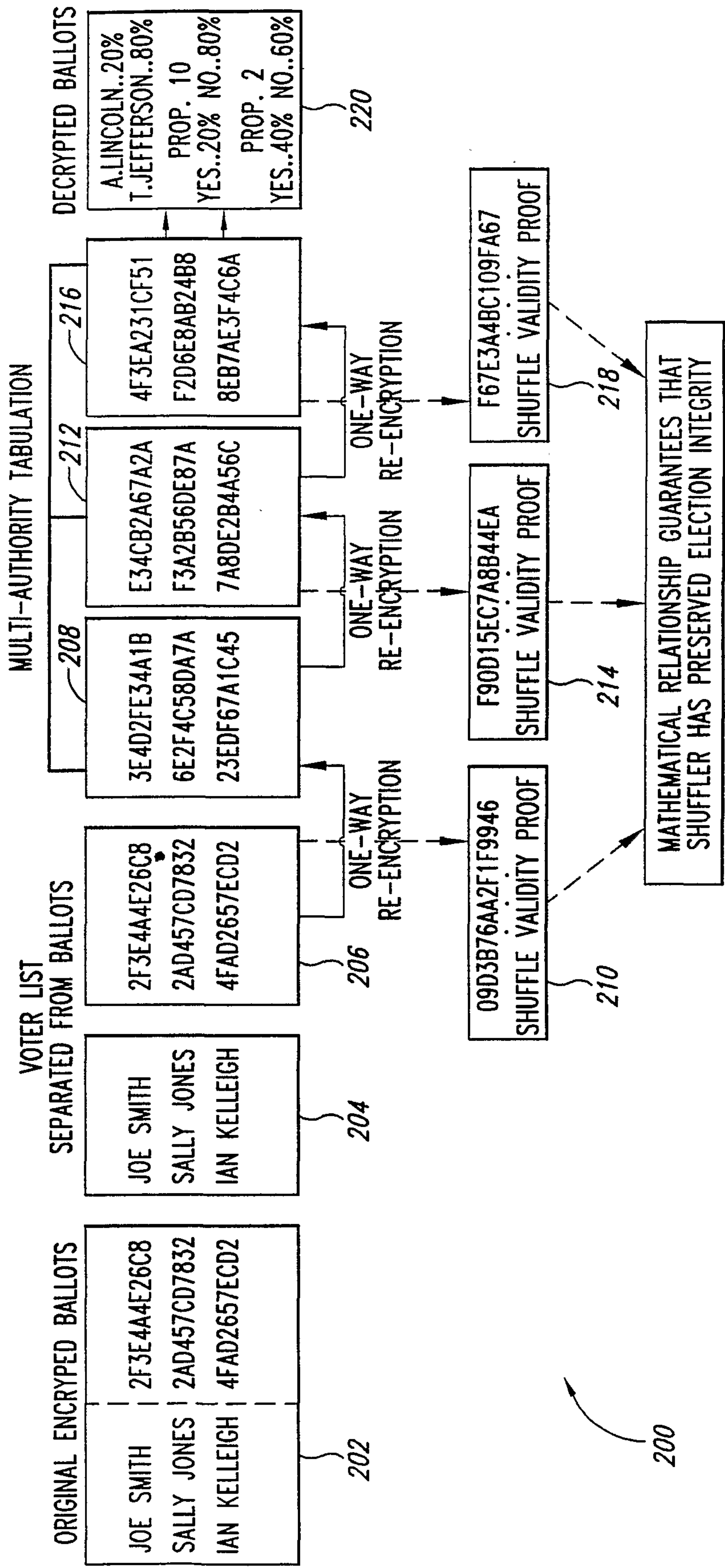


Fig. 2

