eVoting Security & Privacy in

泛民特首候選人初選
Pan-dem Chief Executive Primary Election

SC Leung

CISSP CISA CBCP

Pan-democratic Chief Executive Primary Election

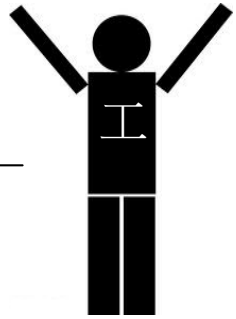
- Objective
  - To allow pan-demo supporters to elect one candidate to join the 2012 Chief Executive Election
    - Let citizens decide who be the candidate, not the parties negotiation
- Format
  - Part 1: e-Voting (8 Jan 2012) 10:00-19:00
    - Citizens voted in 74 voting stations scattered in HK Island (19), Kowloon (24) & New Territories (31)
  - Part 2: Polling (3-6 Jan 2012)
- Result Evaluation
  - Part 1 total valid vote percentage 50%
  - Part 2 total valid questionnaire 50%

The e-Voting Process

Voter

Voting station

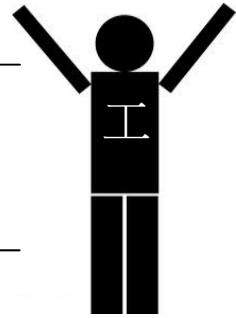
Station staff briefed
Privacy Statement
& Voting instruction



Valid HKID Card

Identity of citizen
Matches HKID

Age >= 18 years



Check if HKID been
used in voting

Choose one option
in person

Success message

Voter returns tablet
to staff

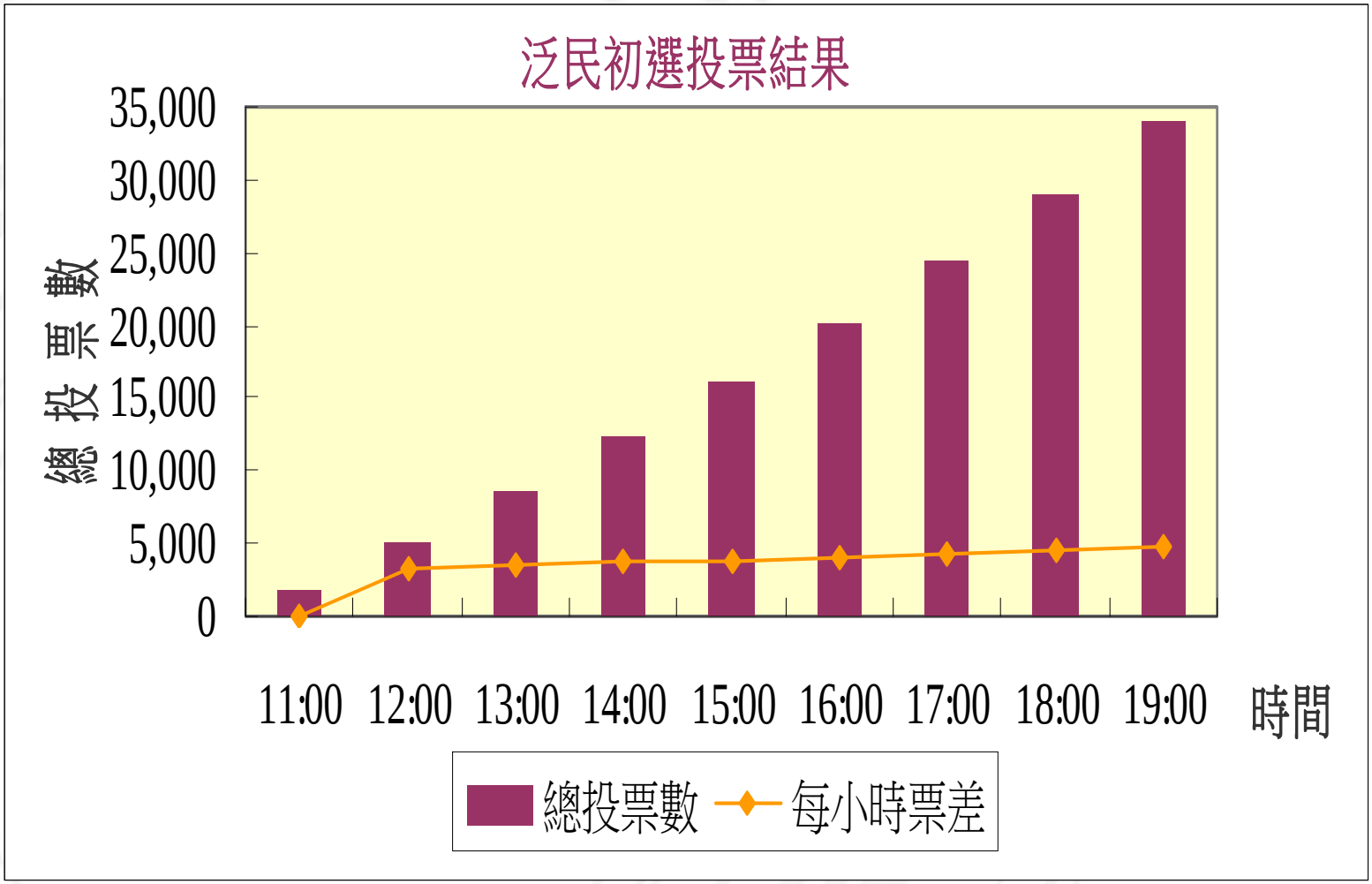
Personnel in the e-Voting Oganization

- Candidates & Representatives
- Voters

- Voting Organization Committee (Pan Demo camp)
- Voting Station Staff (HKUPOP)

- System Developers
- Production Support
- Security Consultant
- Reviewer (and external independent review panel)

Result



http://pdce-primary.hk

Voting Result

Time	Total Votes	Delta
11:00	**1,755**	**0**
12:00	**5,150**	**3,395**
13:00	**8,572**	**3,422**
14:00	**12,237**	**3,665**
15:00	**16,063**	**3,826**
16:00	**20,157**	**4,094**
17:00	**24,418**	**4,261**
18:00	**29,031**	**4,613**
19:00	**33,932**	**4,901**

- 何俊仁 (Albert Ho) 22,148 (67.24%)
- 馮檢基 (Frederick Fung) 10,791 (32.76%)
- 棄權 (Abstain) 993 (2.93%)

Timeline

- Timeline
  - 11 Dec 2011 — Election Committee Election
  - **21 Dec 2011** — **First Project Meeting discussion**
    - Design → development
    - Risk assessment → security design → mitigation
    - Invite panel to review security design
  - 21 Dec 2011 – 05 Jan 2012 — Programming of System Version 1
  - 03 Jan 2012 — Voting Station Staff Training
  - 01-04 Jan 2012 — Third Party Review of Approach
  - 05-06 Jan 2012 — Program amendment, Test and Bug Fix
  - 07 Jan 2012 — Deployment
  - **08 Jan 2012** — **PD-CE D-Day**
  - 09 Jan 2012 — Clear up data

D-Day Operation at Command Centre

- 07 Jan 2012
  - SMS to staff users credential and specific URL
  - Staff scan their mobile devices for malware
- 08 Jan 2012
  - 09:00 Developer, Production Support, Observer gather at Command Centre with Organization Committee
  - 09:55 Zero of ballot box
  - 10:00 Start Voting System
    - Hourly report of voter counts
    - Production support (reset locked accounts and security incident reponse)
    - Program support, if required, under supervision of Production Support and Observer
  - 18:30 SMS to staff station close at 19:00 and cut the queue
  - 19:30 All stations closed
  - 20:30 Open votes in front of candidates (Representative), Voting OC and Auditor
  - 21:00 Clear hash data on servers
- 09 Jan 2012
  - Overwrite server hard disk with data erasure tool

Risk Assessment of e-Voting Process

- *Availability*. The availability of the system must be ensured during the primary election period
- *Fairness*. Duplicated vote has to be avoided
- *Secret Ballot and Confidentiality*. Breach of voting decision and correlation with personal identity must be avoided. Distribution of vote should be kept confidential before opening of votes.
- *Privacy*. Personal identifiable information used in the process must be protected.

Risk Assessment of e-Voting Process

- *Identity*. Because e-voting is conducted over the Internet, the system must prevent spoofed e-voting portal or spoofed voter.
  - Spoofing website can be used to conduct phishing or man-in-the-middle attack.
  - Spoofed voter can damage the integrity of voting data.

- *Integrity*. The integrity of data must be protected from hackers or malicious insiders.
  - The portal server must be protected from hacking attacks.
  - The voting devices, being brought about by station staff, must be protected from contamination.

Risk Mitigation Measures

Availability

- The e-voting system is hosted in a secured Internet data centre.
- The front-end is protected by CloudFlare cloud security service that provides
  - blocking or access by geographical location to avoid non-Hong Kong network access, thus limiting overseas network attacks and greatly mitigate DDoS attacks.
  - block web application attacks
- A backup server is available on Amazon cloud to provide resilience to the system against DDoS attack and high visitor traffic
  - Data is synchronized to backup server in real time. Alternate voting portal is available for people to access in case one site is down.

Fairness

- One Voter cannot have duplicated votes.
- Voters are verified physically with their HKID card.
- A unique hash fingerprint value derived from a proprietary function including the voter's HKID number is stored on the database to signify a person has previously voted. The voting system bans duplicated voting.
- Voting result only disclosed after the voting is closed.

Secret Ballot and Confidentiality

- The personal HK identity number of the individual is only used to validate the person at the polling station and checked against multiple voting attempts in the system.
- There is **no correlation** between the **hash fingerprint** generated from the voter's HKID number and the **voting decision**. Furthermore, the two sets of data are stored in separate tables.
- The **transmission** of voter decision to e-voting system is authenticated and encrypted by **HTTPS**, an international standard for authentication and encrypted data transmission. The database synchronization between production server and backup server is on encrypted channel.

Privacy Protection

- Besides the HKID number, no other personal identifiable information is collected.
- **The HIID # hashed** with strong encryption algorithm to protect its real value to be exposed or even stored in the system.
- The web portal is equipped with virtual keyboard to avoid any keylogging software from logging credentials entered on physical keyboard.
- The e-voting system is a dedicated and an independent system. It is protected by a set of independent strong password.
- **Voter decision and voting history are stored in separate tables on a centralized e-voting system.**
- **No data is stored on the voting device**. Compromise to a voting device cannot grab any data of the voter from the local device.

Hash Function

- Hash = a fixed length bit string that represent a text block uniquely
  - Hash algorithm is mathematical hard problem:
    - easy to generate the hash of a text block
    - but (computationally) infeasible to recover the text block from the hash
  - Avalanche effect: change of a bit in the text block drastically change the hash
  - Collision resistance: hard to find two text block that generates the same hash

- Use of hash
  - Signature → verify integrity of file/data
  - Generate key for password to be stored (resist guessing by administrator)

Attacks to Hash

- Weak hash algorithm
  - If collisions is found, hash algorithm strength degrades
    - MD4 (128b), MD5(128b), SHA1(160) were found collisions
  - Currently SHA2 (SHA-256. SHA-512) is not yet broken
- Brute force attack
- Rainbow table attack
- Challenge to Hashing of HKID #
  - HKID# has a relatively small search space (only 8 characters/digits)
    - $26^2 * 10^6 = 6.76 \times 10^8$ combinations

General Defense Strategy to Attack of Hash

- Using SALT
  - SALT = an additional string appended to text block to expand the search space
  - The SALT and its length should be kept secret
  - Example
    - If a SALT is added to **make the text block 20 characters long**, and the SALT is taken from ASCII (say 94 characters), then the search space becomes

      $94^{20} = 2.9 \times 10^{39}$ combinations.
    - If a computer can compute 100K SHA512 per second, it takes **9.2×10^{26} years** to do exhaustive search in the space
      - **the universe is said to be between 1.37×10^{10} years old**

- Iterative Hashing
  - Taking a particular hash function and hash it, say 1000 times, to make it harder (1000 times computationally) to crack

Hash Function used by the Pre-Election

- The hash algorithm is based on SHA-512 algorithm, which has no known attack.

- Added layers of defenses
  - (1) Add secret input format to HKID#, e.g. "A1234567 → aA-123-456[7]". Then hashed with SHA512 to form a 128 bit string
  - (2) SALT: variable length (5-20 characters) depending on HKID# hashed by a public strong algorithm (other than SHA512) to 64 characters before use
  - (3) perform transposition of the hashed HKID# string and the SALT string (128+64 = 192 chars) using a function depending on HKID#
  - (4) variable number of iterations (100-1000) of hashing depending on HKID#

- To reverse the HKID #, an attacker needs to know:
  - the salt table + the program logics of all there variable functions

- Even if attacker succeed in reversing one text block, the same salt and function won't apply to another HKID#

How strong is the hash function against insider attack?

- To protect from insider attack, the code (salt + program logic) is encrypted during compilation
- How strong is the algorithm towards its weakest link -- the developer?
  - The huge search space and variable function would make the developer difficult to generate the lookup table or rainbow table to break his own system.
  - After deployment of system, he is denied access to the system
  - During voting day, he has to stay in the command centre with an Observer and the Production Support team. If program bug fix is required he is under supervision to conduct his work and forced to keep hands off after recompilation of program.
- Another protection against Insider attack
  - Before the deployment, the source code (except SALT & variable hash functions) are reviewed to avoid back door planting.

- Note: after all, breaking this hard problem only allows attacker know that a person has voted, but not the decision of vote, since they are in two tables.

Integrity

- The voting portal server is well protected by firewall and cloud based web protection services. Web application intrusions are filtered. Security patches are updated to the latest suitable level. The URL of the e-voting portal is kept secret and disclosed to voting station staff on the voting day via SMS (alternate channel).
- The e-voting system is verified to be free of backdoor before deployment and after putting up on the production system.
- The voting devices are provided by the voting staff. The voting staff are recruited by HKU and most of them are experienced in public survey alike activity. The voting machines are scanned by security software before taking to use in the voting. All staff must sign a compliance and non-disclosure statement to guarantee they follow the guideline regarding security and privacy.

Identity

- The identity of the e-voting portal is authenticated by SSL certificate issued by trusted certificate authority. The validity of the certificate is used to identify the identity of the site and to prevent man-in-the-middle attack.
- Each voting staff account is bound to a single voting device by a special token. Logging in from a second device will disable the session and staff account.

Operational Security

- *Separation of Duties*.
  - Voting station staff and supporting staff are isolated in the system in terms of access point and account.
  - Voting staff cannot see the system-monitoring status and system-monitoring staff cannot see the voting interface.
  - All program codes, not only related to the hashing, are encrypted, so that supporting staff has no access to the program logic and any sensitive information inside the code. The encrypted program expires in 36 hours and no longer operable. Re-encryption and re-deployment is needed afterwards.
  - The developer holding the database password has no access to the production system after full deployment. If there is a need for troubleshooting by the developer, he is escorted by an observer nominated by the voting office and his activity is monitored.

Operational Security

- *Incident Response.*
  - A team of supporting staff will be on duty on the voting date, to monitor and response to system health and abnormal user pattern, e.g. in a short period of time, input a lot of record to the system. Supporting staff can revoke the access of any voting staff to prevent suspicious activities, or reset passwords. Password reset requests would be investigated and verified and sent via SMS or phone number previously determined.

Authentication and Access Control

- Each voting staff has a unique userID, password and a unique obfuscated URL to access the voting portal. Mismatched entries of the three will be rejected for accessing the site.
- The voting station staff accounts, the monitoring staff accounts and the counting staff accounts have strong passwords.
- The server-monitoring page is limited access to pre-defined IP addresses of the voting office.

Personal Data Privacy Protection

- The e-voting system has been reviewed under the six Data Protection Principles issued by the Privacy Commissioner's Office):
- **Principle 1** -- Purpose and manner of collection
  - The voter is explained, before collection of the HKID number, of the purpose of collecting their HKID number, which is to avoid duplicate voting. The HKID number is collected in person at the poling station by official voting station staff. After successful voting, the same message of the purpose of collecting their HKID number is again displayed on screen.
- **Principle 2** -- Accuracy and duration of retention
  - The HKID number is hashed right after entering into the system to protect its true value as well as ensuring its accuracy. Any chances of alteration will be identified by the system and security administrator.
- **Principle 3** -- Use of personal data
  - The HKID number will only be used in identifying duplicate voting. No other personal identifier will be collected and usage is allowed.

Personal Data Privacy Protection

- **Principle 4** -- Security of personal data
  - In addition to the encryption of the HKID number by the use of digital hash function of SHA-512 algorithm, the system is also protected by network security devices and gone through security tests before production. For details, please see “Information Security Protection” section above.
- **Principle 5** -- Information to be generally available
  - Since the HKID number is never really going to be stored in the system, other than its hash value, there is no need to make it available for the voter to access and verify for its accuracy.
- **Principle 6** -- Access to personal data
  - Same as above. Also, the production database system (the non-cloud one) will be digitally cleaned using DOD_5220.22-M compliant methodology after the completion of the voting process.

Improvement wanted

- Trade-off: *Anonymity vs. Auditability of voting decision*
- More More More time please
- Audit the implementation, not just the design, if more time allowed
- SALT can be generated by the person other then the developer in the beginning of voting.
- Integrity improvement:
  - Use PKI to sign every vote. The keypair be generated in front of auditor, vote O.C. and observer. Private key being held by auditor, and used to decrypt every vote during vote counting.
- More …

Q & A