Mixing Email with BABEL’96

Ceki Gulcu, Gene Tsudik
IBM Zurich, Switzerland

cs6461, Fall 2008
Computer Science, Michigan Tech
Byung Choi
Introduction

• The most critical network building block for anonymity: Mix by Chaum’81

• Absolute anonymity: dining cryptographers algorithm and protocol by Chaum’88
  – Impractical due to the large protocol overhead
  – Hard to provide secure pairwise channels and a synchronized broadcast channel

• Advancement being sought
Related work

• Anonymous Email systems
  – Penet: Finland 1990s
  – Cypherpunk: 1990s – now?
  – Mixmaster: 1990s – now?
  – Implementations of the concept of Mix by a single individual or a group of contributors

• Systematic improvements needed
Desired properties (requirements)

- Email systems accommodate anonymity
- Hard to determine the sender
- Recipient can reply the anonymous sender
- End-point anonymity preserved, intermediate mixes are not maximally trusted
- Infrastructure resistant to active attacks
- Sender gets an anonymous confirmation
- Low noise
Notation

\( M \) message; sequence of ASCII bits
\( E_x(M) \) encryption of M with X’s public key
\( D_x(M) \) decryption of M with X’s private key
\( K \{ M \} \) conventional encryption of M with key K
\( (M_1, M_2) \) concatenation of \( M_1 \) and \( M_2 \)
\( A_x \) X’s email address.
\( [M]^{\Omega} \) padding string M to length \( \Omega \)
(by appending random bits)
\( [M]^{\Omega} \) trimming string M to length \( \Omega \)
(by removing trailing bits)
Mix, revisited

Alice \rightarrow Mix \rightarrow Bob

Eve
Attacks

- Passive
  - Content correlation:
    - uniform length, padding, nonce
  - Time correlation:
    - Regular vs. interval batching

- Active
  - Isolate and Identify
  - Message replay:
    - time stamp, message identifier
  - Cascading or chaining mixes
BABEL

• Forward path
  – Composition by sender
  – Processing by mixes
  – What does a mix know?
Forward message

- Encryption with F1’s key
- Other encryptions
- Encryption with F(f-1)’s key
- Encryption with Ff’s key
Forward message

\[ x_f = E_{F_1} \left( A_{F_2}, E_{F_2} (\ldots E_{F_{f-1}} (A_{F_f}, E_{F_f} (A_{Bob}, [M]^\Omega)) \ldots) \right) \]
BABEL

- Return path
  - Creating RPI
  - Replying by recipient
  - Reply processing by BABEL
  - Handling replies at the originator
  - Two-way anonymous conversation
  - Security of replies
  - Inter-mix Detours
  - Indirect replies
$$y_r = A_{R_1}, E_{R_1}(K_1, A_{R_2}, E_{R_2}(K_2, \ldots \ldots E_{R_r}(K_r, A_{Alice}, E_{Alice}(KS, r)) \ldots))$$
Return path information

\[ \omega \text{ [bytes]} \]

- **K1, Address of R2**
  - Encryption with mix R1’s key

- **Kr, Address of Alice**
  - Encryption with mix Rr’s key
  - Other encryptions

- **KS, r**
  - Encryption with Alice’s key

- **Padding**
Reply message

Email (SMTP) Header

RPI

Message Body
Two-way anonymity
Security of replies
Inter-mix detours
Keeping message size constant

Ω bytes

Data

Padding

Encryption

Data

Padding

Trimming

Data

Padding

excess bytes
Heeding anonymity

- Fixed-path systems
- System staunchness, miss & guess factors
- Quest for confusion
  - Probabilistic deferment
  - Hybrid approach
Trickle attack

single message per period

M1 → M2 → ... → Mm

Controlled by Eve

Eve
Interval batching
Probabilistic deferment

\[ P\{K = k\} = \binom{m}{k} q^{m-k} d^k \text{ where } k = 0, \ldots, m, \]
Binomial function

Odd number of mixes

Even number of mixes

Probability of the most likely b vs. Deferment probability, d.