COST TO DEFEAT THE N-1 ATTACK

Neil Chettiar

The N-1 Attack

- Different Variations:
  - *Threshold Mixes*
  - *Timed Mixes*
  - *Pool Mixes*

- Attacks are Exact and Certain.
- Basic idea: *Attacker* should be able to account for all the outgoing messages, save for one.

Approach

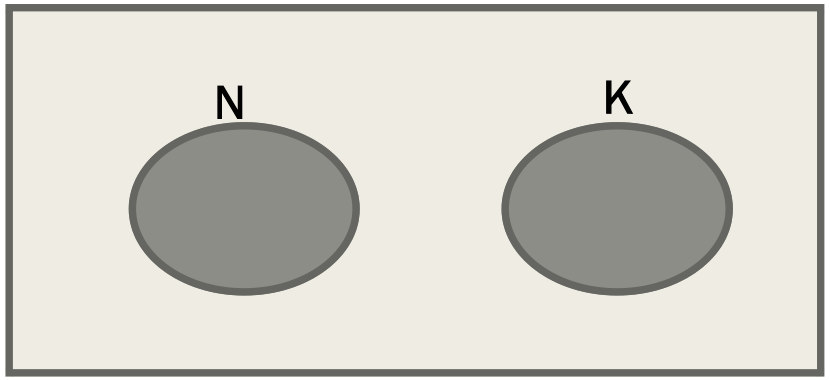
- Use dummy messages

- This still fails for:
  - *Threshold*
  - *Timed*
  - *Threshold AND/OR Timed*

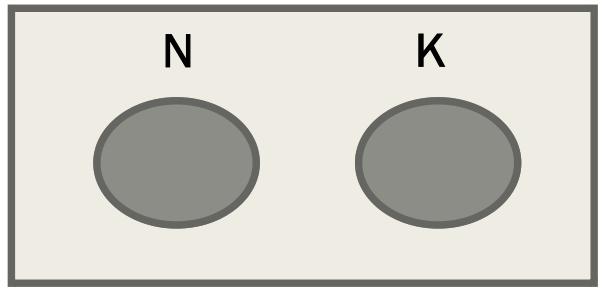
- Attack is inexact but certain.
- The attacker has high probability of determining the trigger conditions

Proposed Mix

- Pool mixes offer a better anonymity metric



Proposed Mix

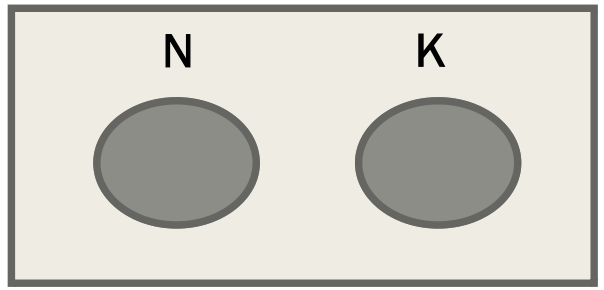


- Selecting the genuine messages:

$$n_i = \begin{cases} 0 \text{ if } N_i < N_0 \\ [N_0, N_i/2] \text{ if } N_i > N_0 \text{ and } N_i/2 > N_0] \\ N_0 \text{ if } N_i > N_0 \end{cases}$$

- K_i = RAND[n_i /2, n_i]

Proposed Mix

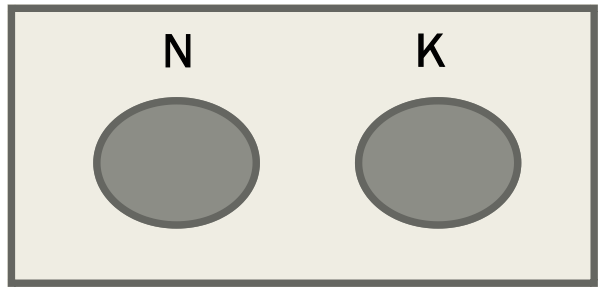


- Probability of flushing the input message from the mix:

$$p(round_i) = \begin{cases} p(n_i) \text{ if } r = i \\ p(n_i) \prod_{j=r}^{i-1} (1 - p(n_j)) \text{ if } r < i \end{cases}$$

$$p(n_i) = \frac{n_i}{N_i}$$

Proposed Mix

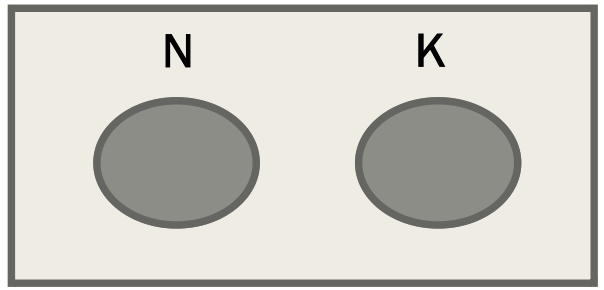


- Anonymity set size:
  - *Def: Probability of linking the output message to the input message.*

$$p(I_i) = \begin{cases} \frac{1}{n_i} \text{ if } i = r \\ \frac{1}{n_i} \prod_{j=r}^{i-1} (1 - p(n_j)) \text{ if } i < r \end{cases}$$

  - *Without considering Dummy traffic.*

Proposed Mix

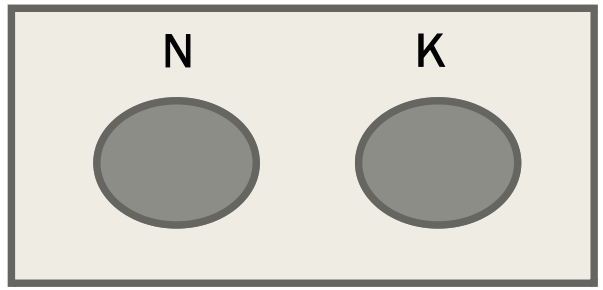


- Entropy:

$$H_s = -\sum_{i=1}^{r} a_i p(I_i) log_2(p(I_i))$$

  - a_i *:Is defined as the number of messages that arrive in the ith round.*

Proposed Mix



- Entropy with Dummy traffic:

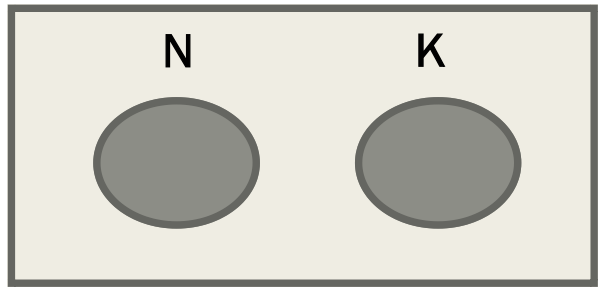
$$p_d = \frac{k_i}{n_i + k_i}$$

$$p(I_i) = (1 - p_d) * p_i$$

$$p_i = \frac{p(n_i) + k_i}{a_i}$$

  - *p_i is defined as the probability of the message being chosen from the input.*
    - *Source: Ref. 1*

Proposed Mix



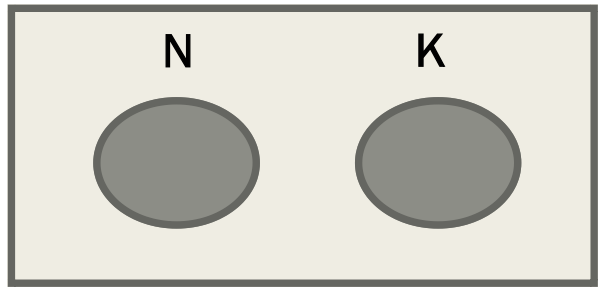
- Entropy with Dummy traffic:

$$H = -p_d log_2(p_d) - \sum_{i=1}(1 - p_d)p_i log_2((1 - p_d)p_i))$$

  - *Entropy for the dummy messages can be calculated separately and added to the entropy for the genuine messages.*
  - *Uncertainty of dummy messages decreases when $(1 - p_d) << p_d$*

  - *Source: Ref. 1*

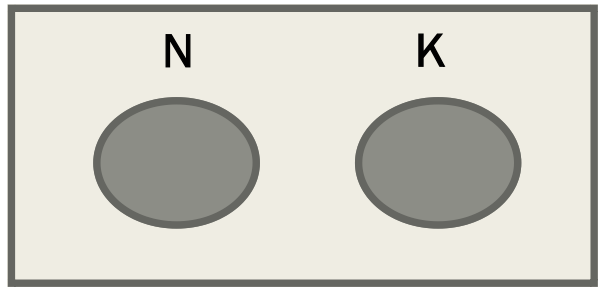
Proposed Mix



- Traffic generation of the MIX:
  - *Probability of finding the target message from the output is:*

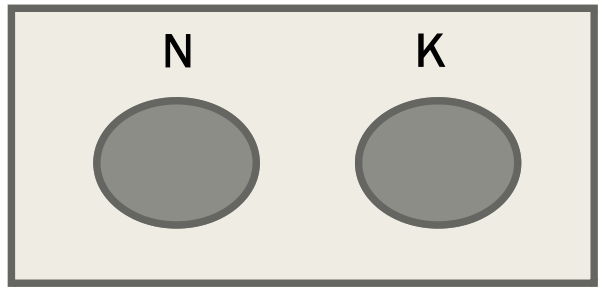
$$= \begin{cases} \frac{1}{n_i+k_i} \text{ if } i = r \\ \frac{1}{n_i+k_i} \prod_{j=r}^{i-1} (1 - p(n_j)) \text{ if } i < r \end{cases}$$

Proposed Mix



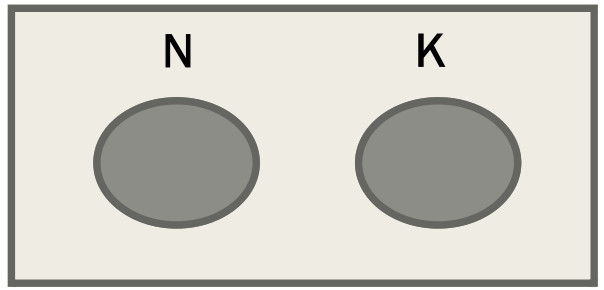
- Simulating the N-1 Attack:
  - *Worst case*
    - *assume* $k_i = n_i / 2$, $n_i = N_0$
  - *The proposed pool mix performs almost as bad as a general threshold mix with dummy traffic, attack is inexact and uncertain.*
  - *The work of the attacker exponentially increases*
  - *This model shows a blending attack anonymity k.*

Proposed Mix



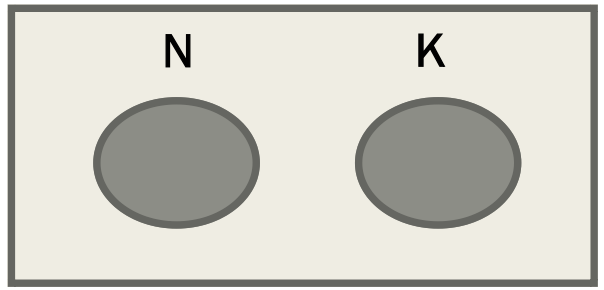
- Counter Measures to the N-1 Attack:
  - *Heartbeat traffic*
    - *Use some of the dummy traffic to check if an attack is in progress.*
      - *Wait till the attacker is exhausted*
      - *Drop all the messages 'Push the Red button'*
  - *Sending crisis messages to the other mixes in the network*
    - *The mixes receiving the crisis message can drop the targeted mix from its mailing list and drop all messages from it.*
  - *Regroup-and-Go Mixes*
    - *The sender can segment his message before sending, attacker cannot collect all the information.*

Proposed Mix



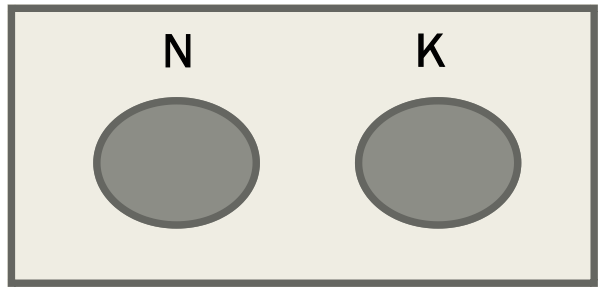
- Dummy Traffic Hop Count.
  - *For cases when the Hop count is >1, the problem can be translated to an equivalent problem of adding dummy messages into the pool*
  - $N_i = a_n + a_k$
  - *Probability of the genuine message being chosen is:* $n_i / (a_n + a_k)$
  - *=> chances of selecting the genuine message decreases.*

Proposed Mix



- Cost in terms of traffic generated:
  - *Difficult to model because of the different types of connections and bandwidths that exist in real world scenarios.*
  - *Simulating a small scale system, would not really help because of the limitations on network and bandwidth.*
  - *Load estimation techniques*

Improvements to the Proposed Mix - Conclusions



- Changing the threshold according to activity.
  - *If the mix is inactive over a long period of time, lower the threshold.*

- Improvements to the counter measures:
  - *For cases when the heartbeat traffic could be wrongly assessed:*
    - *Shift the load to other mixes [TOR's Loose Routing]*

References

1. Reasoning about the anonymity provided by pool mixes that generate dummy traffic. – *Claudia Diaz and Bart Preneel*
2. Heartbeat traffic to counter (n-1) attacks. - *George Danezis and Len Sassaman*
3. From a trickle to a Flood: Active attacks on several Mix types. - *Andrei Serjantov and Roger Dingledine and Paul Syverson*
4. Regroup-And-Go mixes to counter the (n-1) attack. - *Jin-Qiao Shi, Bin-Xing Fang and Li-Jie Shao*
5. On blending attacks for mixes with Memory extended version. - *Luke O'Connor*
6. Active attacks on Reputable mix Networks. - *Longhai Li, Shaofeng Fu, Xiangquan Che*
7. Untraceable electronic mail, return addresses and digital pseudonyms. - *David L. Chaum*