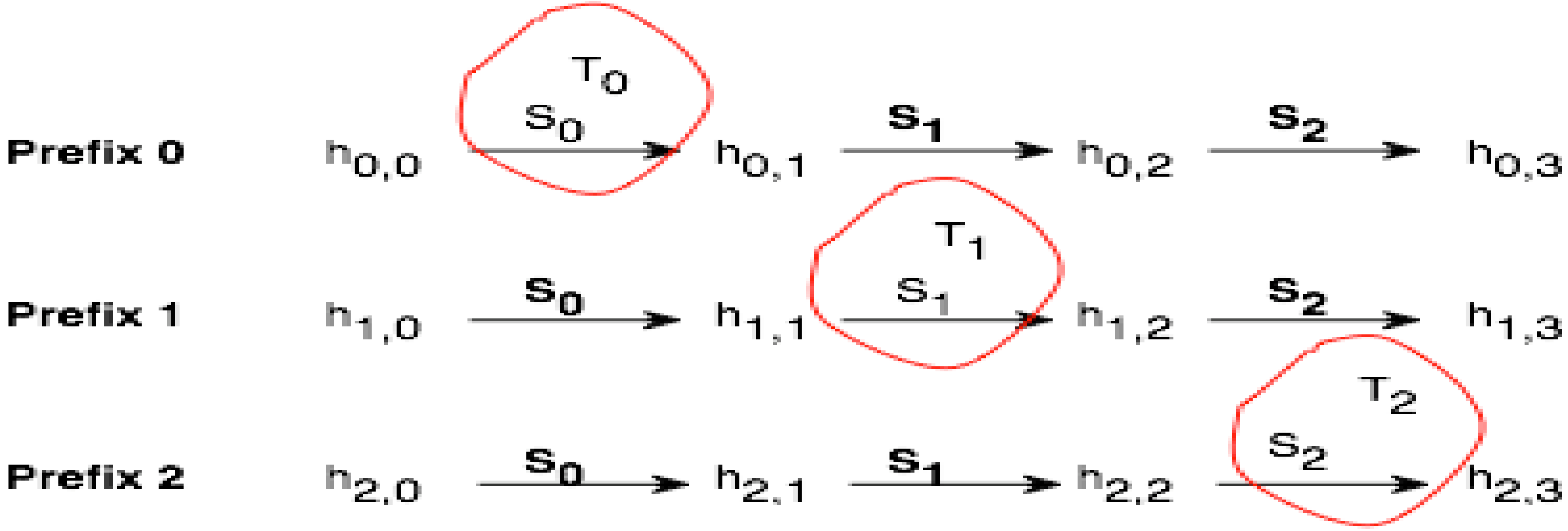
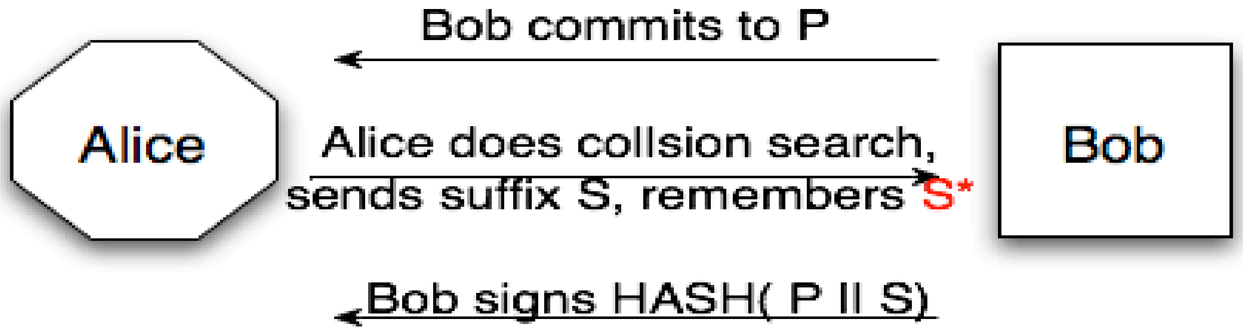
Trojan Messages: Collisions with an unknown prefix



Elena Andreeva, Charles Bouillaguet, Orr Dunkelmann, Pierre-Alain Fouque, Jonathan J. Hoch, **John Kelsey**, Adi Shamir, and Sebastien Zimmer

Alice can break our hash….

UNSAFE: Alice commits to suffix after she knows prefix



Bob commits to P

Alice

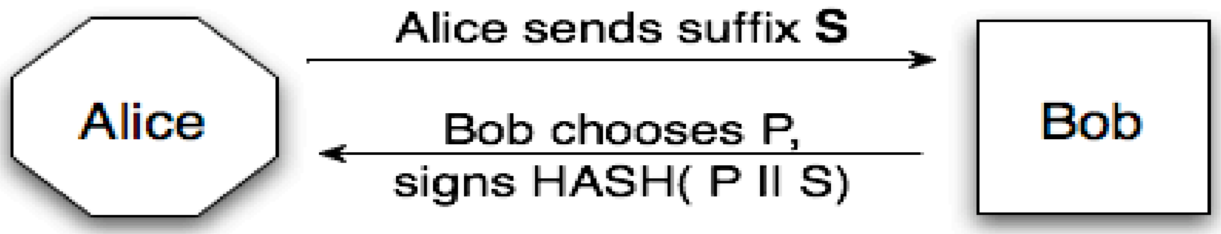
Alice does collsion search, sends suffix S, remembers S*

Bob

Bob signs HASH(P II S)

Alice CAN cheat Bob--she knows collision for HASH(P II S)

SAFE: Alice commits to suffix before she knows prefix



Alice sends suffix **S**

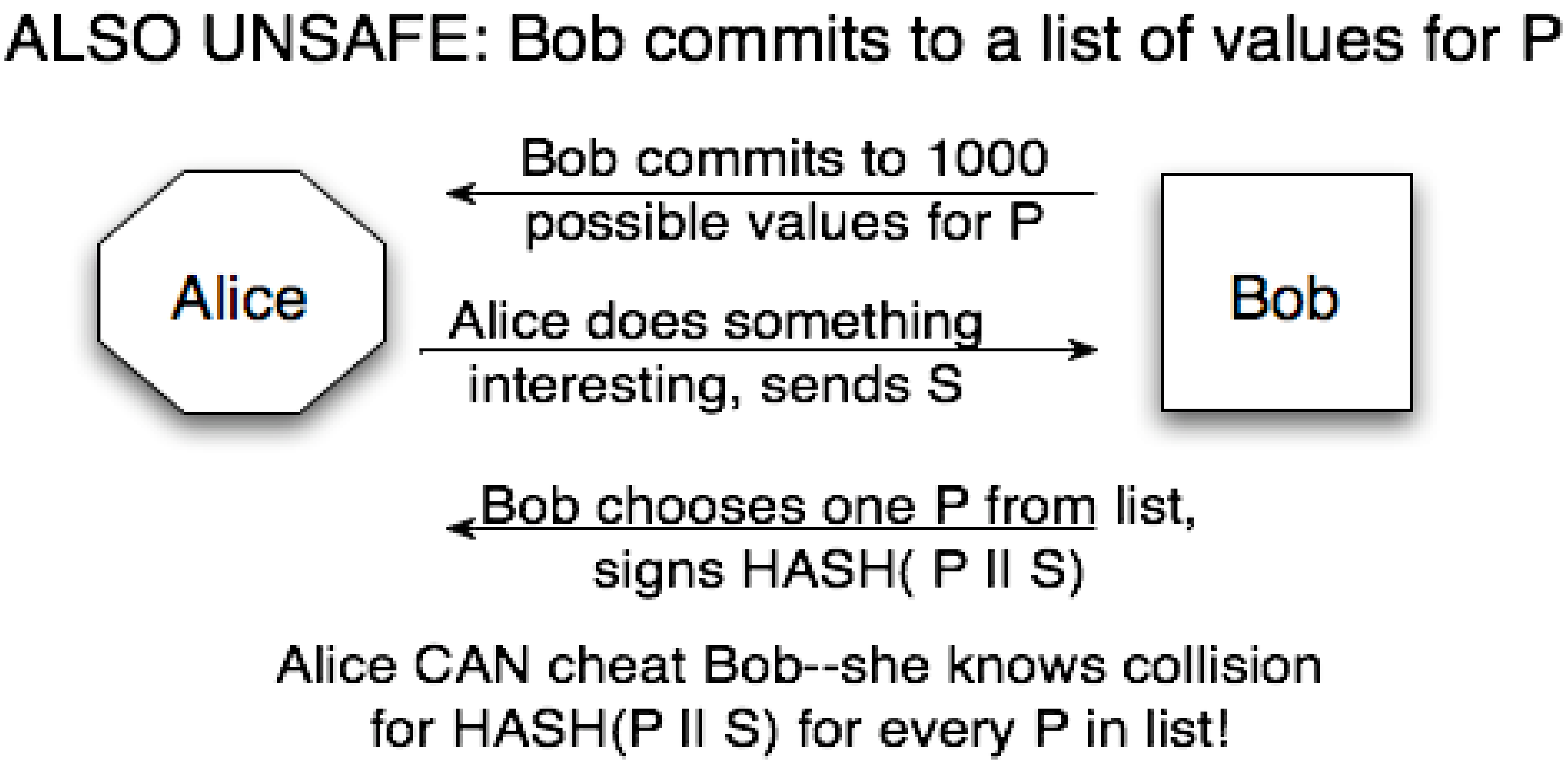
Alice

Bob chooses P, signs HASH(P II S)

Bob

Alice can't cheat Bob--she doesn't know collisions for HASH(P II S)

This Result



ALSO UNSAFE: Bob commits to a list of values for P

Bob commits to 1000 possible values for P

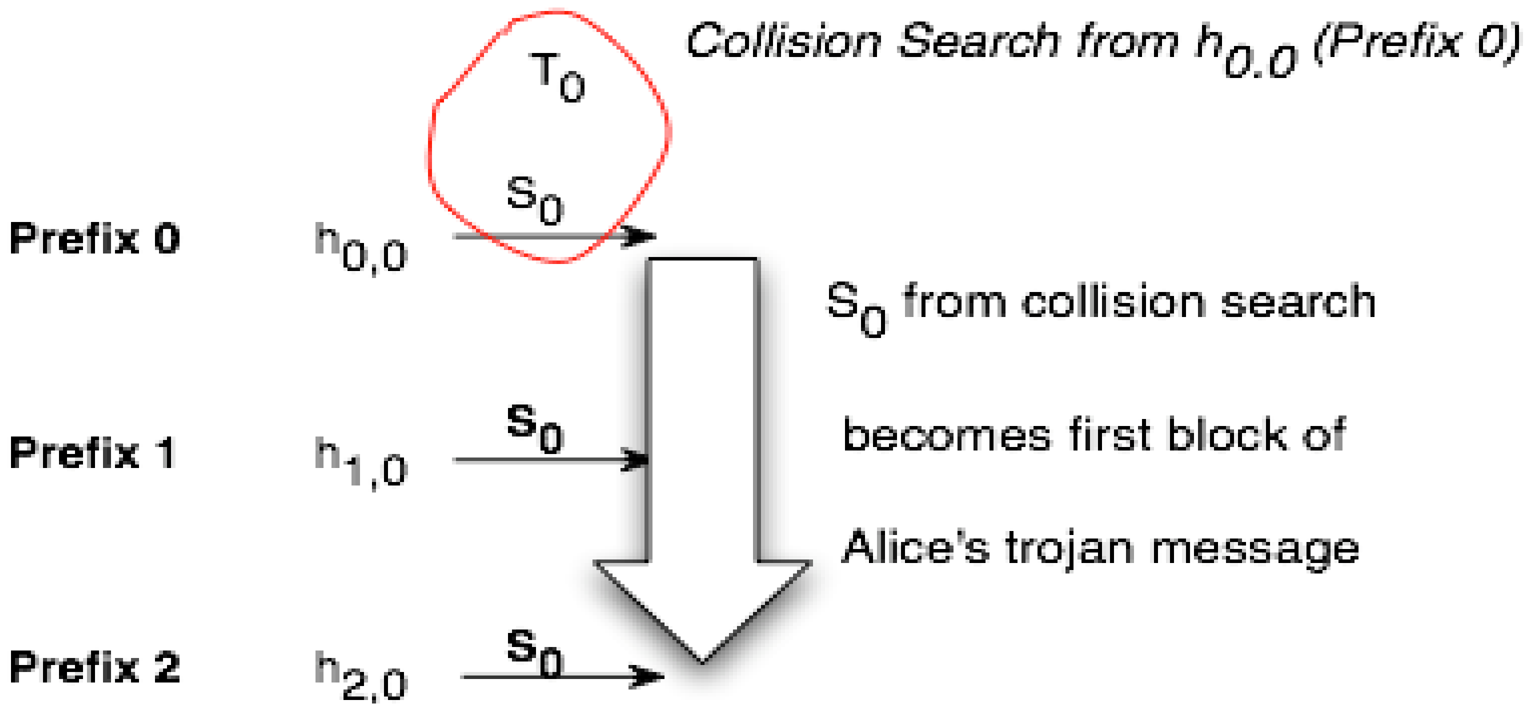
Alice does something interesting, sends S

Bob chooses one P from list, signs HASH(P II S)

Alice CAN cheat Bob--she knows collision for HASH(P II S) for every P in list!

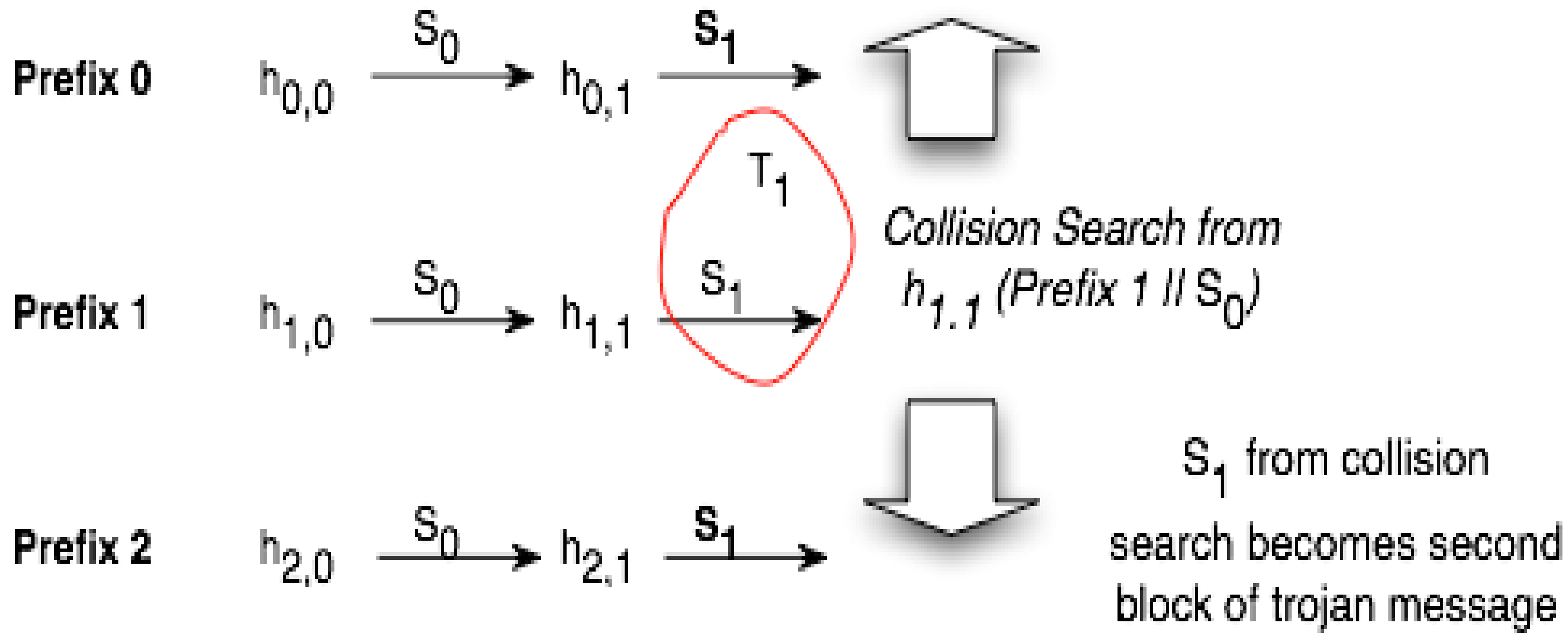
Scales linearly: N possible prefixes -->

N block suffix, requiring N collision searches



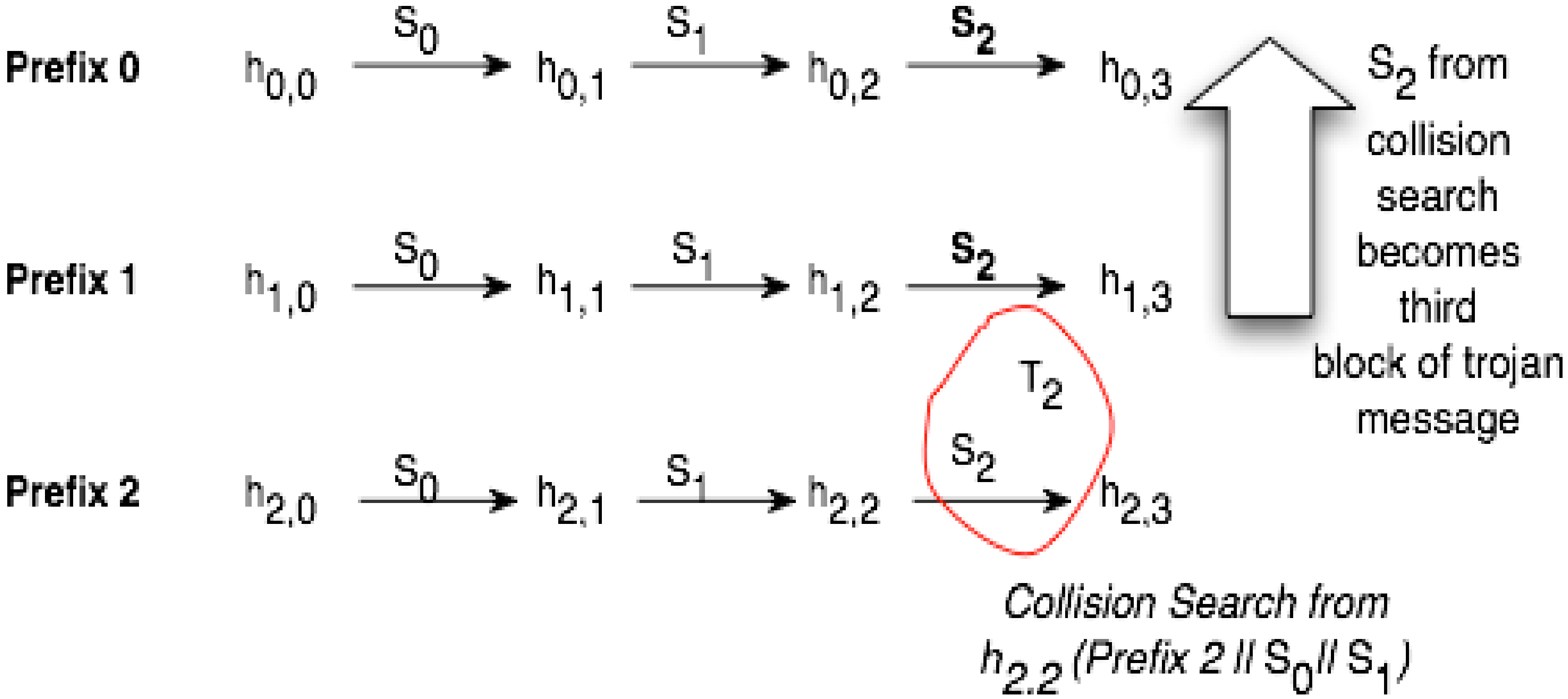
- Do first collision search from Prefix 0
  - (starting from $h_{0,0}$)
  - Search yields S_0,T_0
  - S[0] becomes first block of Trojan message
  - Alice remembers T_0 in case Bob uses Prefix 0

Trojan message so far: ***S_0*****.**



- Do next collision search from Prefix 1
  - (starting from $h_{1,1}$)
  - Search yields S_1,T_1
  - S_1 becomes next block of Trojan message
  - Alice remembers T_1 in case Bob uses Prefix 1

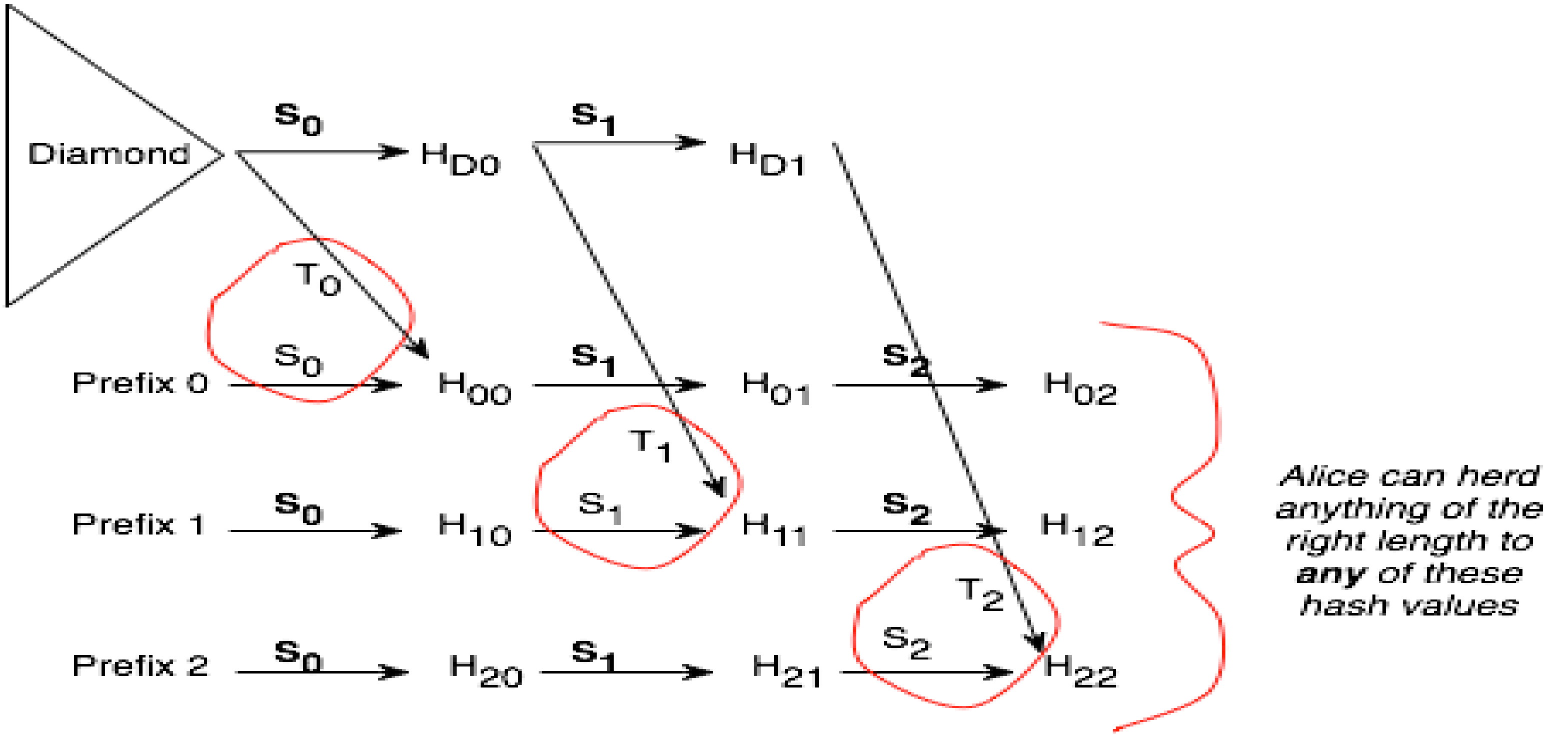
Trojan message so far: $S_0 \| S_1$

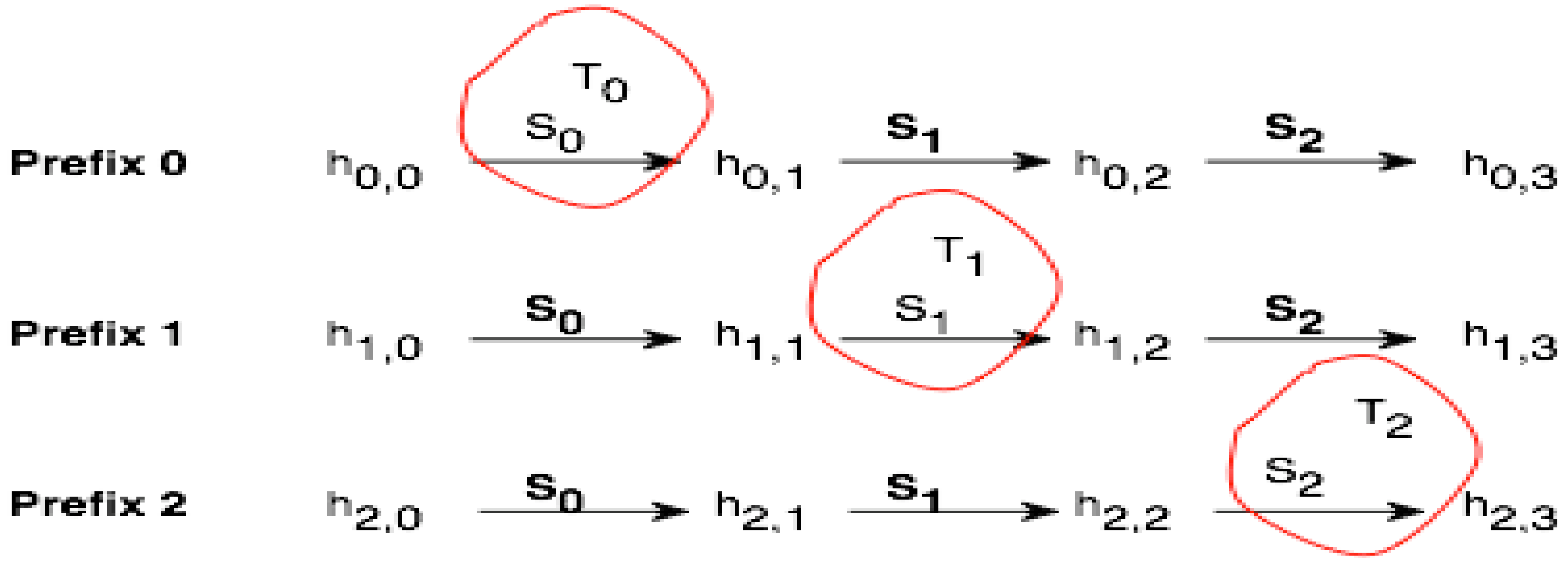


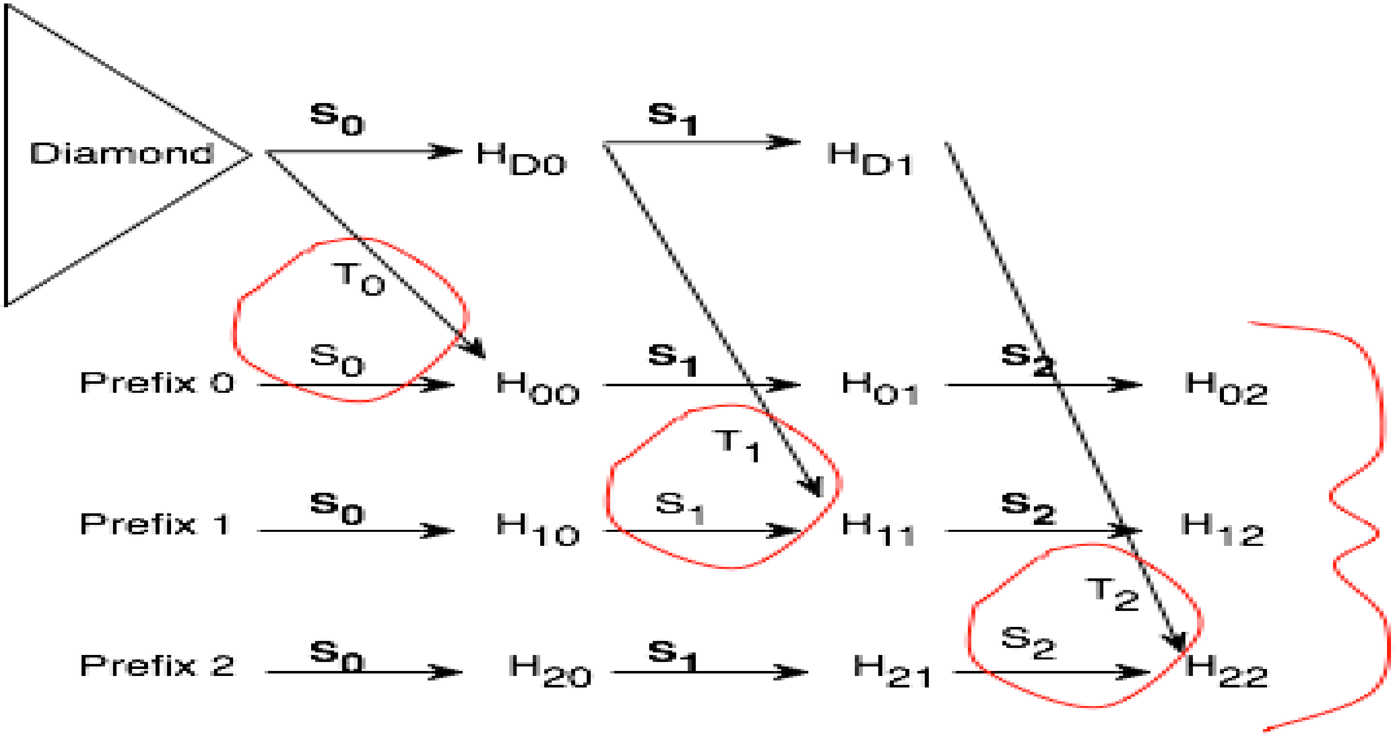
- Next collision search:
  - Starting from **Prefix 2|| S_0 || S_1** (starting from $h_{2,2}$)
  - Search yields S_2,T_2
  - S_2 becomes final block of Trojan message
  - Alice remembers T_2 in case Bob uses Prefix 2

Final Trojan message: S = S_0 || S_1 || S_2

- Herding Variant
  - – Bob commits to N prefixes
  - – Alice builds diamond, sends S
  - – Bob chooses P
  - – Alice can herd any prefix to HASH(P || S)







Alice can herd anything of the right length to ***any*** *of these hash values*