Logical Cryptanalysis:
New Techniques and Future Directions

•••

Alexander Scheel
Iowa State University of Science and Technology

Quick Bio

Hi, I'm Alex.

- Grew up in Rochester, MN
- Attend Iowa State University
- Honors Research is on Cryptographic Hash Functions
- Part of ISEAGE lab, hosting 5 Cyber Defense Competitions a year.
- Participated in ACM's programming competition.

@cipherboy

alexander.m.scheel@gmail.com

Presentation Overview

What is 3-CNF-SAT?

What's a hash function?

How to apply SAT to Cryptography?

Why is it important?

What is SAT?

SAT is the Boolean Satisfiability problem.

Theoretical:

- NP Complete: Nondeterministic Polynomial Time (verify a solution in polynomial time, finding such a solution takes
- Among the hardest class of decidable problems in CS

Practical:

- History as basis for formal methods, part of proof systems, model checking, etc
- Once thought completely intractable, past decade has shown not only is it possible, but also practical in many cases
- Annual SAT solver competition (usually) in Australia

What's a Cryptographic Hash Function?

A hash function...

- is a deterministic function.
- can "hash" variable length data, but produces a fixed length output.
- is easy to compute, but hard to invert.
- should produce large changes in output with small changes in input.

A cryptographic hash function has...

- *Preimage resistance:* given an output state, the finding a corresponding input should be computationally infeasible.
- *Second preimage resistance:* given an input, finding a second input which hashes to the same output should be computationally infeasible.
- *Collision resistance:* finding any two inputs which hash to the same output should be computationally infeasible.

Uses of Cryptographic Hash Functions

- Signing Documents (RSA, ECC, etc.)
- TLS as a Pseudo-Random Function (key derivation)
- Hash-based Message Authentication Code (HMAC) construction for message integrity
- Unique identifier of a file (e.g., AWS).

Applications to Cryptographic Hash Functions

Measuring the Utility of Collision Classes

Analyzing the Structure of the Collision Class Space

Finding New Collision Classes

Five Utility Metrics

1. The number of unique differentials a collision class has.
2. The number of unit distance neighbors a collision class has.
3. The maximum count of zeros in the binary of a colliding input block.
4. Whether there exists a block which collides under multiple initial values.
5. Whether zero, one, or both blocks may be ASCII valued.

Tables

Number of Differentials for Existing Collision Classes			
Attack	Size	Attack	Size
X. Wang's	64	Y. Sasaki's	4
H. Dobbertin's	32	P. Kasselman's	32
M. Schlaffer's	64		

Number of Zeros in a Colliding Block			
Attack	Count	Attack	Count
X. Wang's	509	Y. Sasaki's	494
H. Dobbertin's	504	P. Kasselman's	504
M. Schlaffer's	506		

The Structure of Collision Classes and Families in MD4

- Probabilistic relationship among classes/families ($\lesssim$):

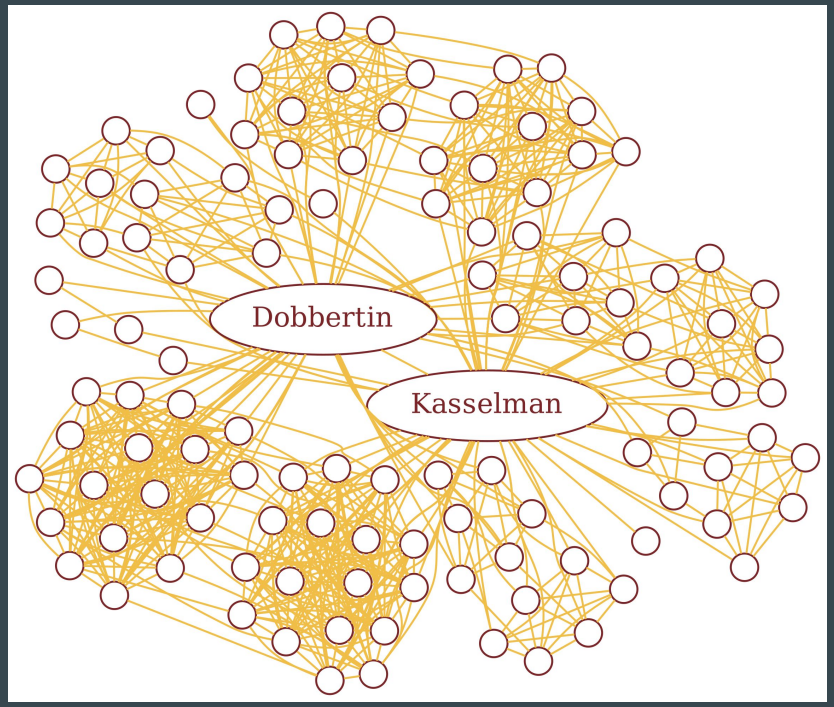
For a given family (F) of collision classes:

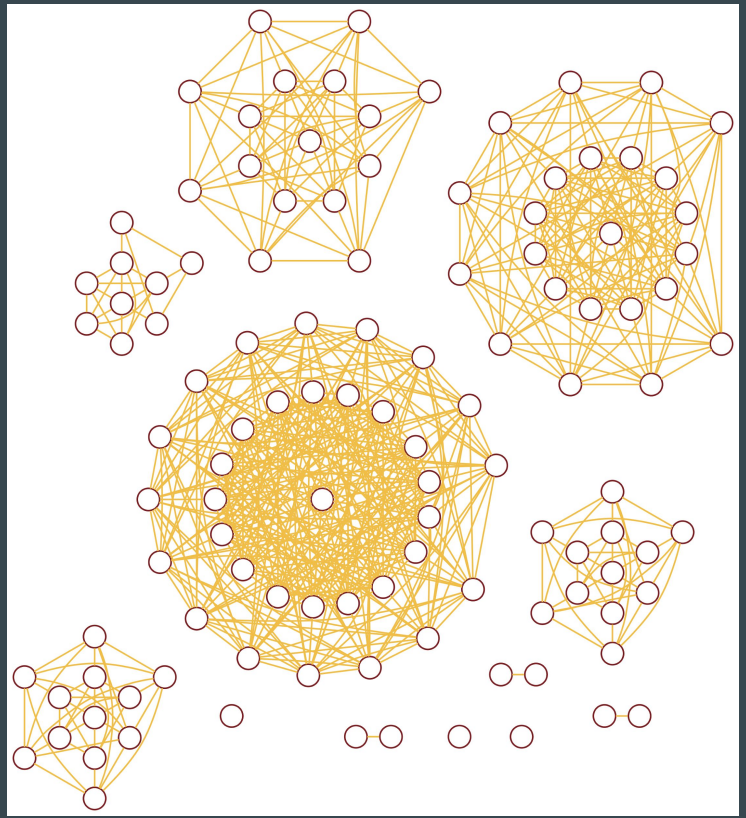
- For every F, there exists G such that $G \lesssim F$.
- For nearly every F, with high probability, there exists H such that $F \lesssim H$.

Where $\lesssim$ acts as a subset operator of families, and operates across rounds.

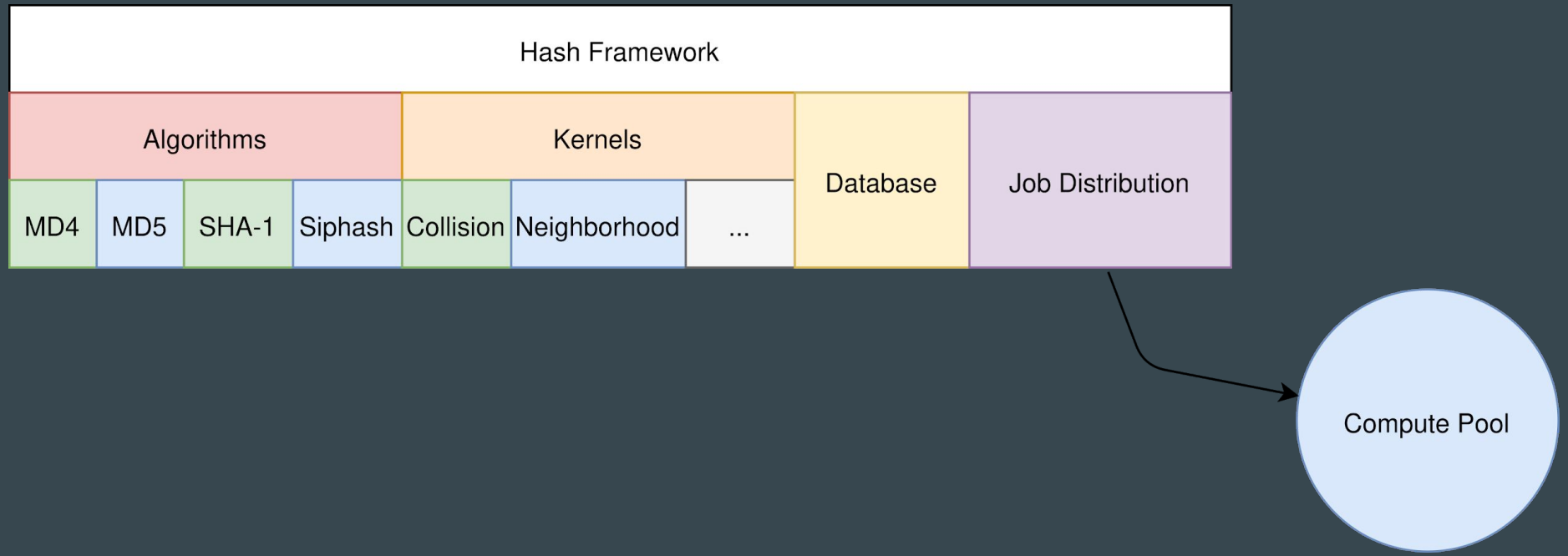
The Structure of Collision Classes and Families in MD4

- Neighborhoods of collision classes.





Exploiting the Structure: 35k New Collisions



Why is this important?

- Utility of collisions.
- Applications of specific techniques.
- Implications for second preimage attacks.
- Non-zero information gain via correct SAT usage.

Questions?

Paper: https://github.com/cipherboy/papers/raw/dsn-2018/dsn-2018/hash_dsn_2018-candidate-2.pdf
Framework: https://github.com/cipherboy/hash_framework