CALIFORNIA STATE SCIENCE FAIR
2007 PROJECT SUMMARY

Name(s)
Gregory A. Hirshman

Project Number
S1307

Project Title
Differential Cryptanalysis of MD5: Unscrambling the Hash Bit by Bit

Abstract

Objectives/Goals
The recent successful attack on the widely used hash function, the MD5 Message Digest Algorithm, was a breakthrough in cryptanalysis. Original papers, published in 2004 and 2005 by Wang and Yu and Liang and Lai, described this attack in an obscure and elliptical manner. Hawkes, Paddon, and Rose later presented the attack in more detail, but even their paper contained numerous unproven statements. This paper will prove assertions made by Hawkes, Paddon, and Rose, provide original corrections and illustrations, and explicate the two primary papers to make them more accessible to the mathematically-literate reader.

Methods/Materials
This paper provides background information on cryptography, hash functions, the MD4 and MD5 Message Digest Algorithms, a substitution-permutation network (SPN), differential cryptanalysis, and the differential attack on MD5 as originally presented. Then, it explicates this attack. Two blocks are treated. For the first block, it adds calculational details, examples, and original proofs to elucidate the step by step analysis presented by Hawkes, Paddon, and Rose. For the second block, it develops a step by step analysis based on a few tables that they provide. Finally, it presents an original comparison between the structures and cryptanalyses of the SPN and MD5 algorithms, providing insight into the attack on MD5.

Results
This paper makes four important contributions. First, it provides an example for each of the three conditions at the beginning of the description of the first block differential, demonstrating why certain conditions had been placed on the Tt. Second, it proves conditions specified by Hawkes, Paddon, and Rose for the first block. Third, it provides for the second block a completely original step by step analysis of both the description of the differential and of the propagation of the differences through the ft functions. Finally, it reveals several mistakes in the work of Hawkes, Paddon, and Rose. Most are trivial, but one is of special importance since implies that their attack is actually about twice as fast as they believed.

Conclusions/Discussion
The goal of this paper is to use original analysis to provide a more detailed, accurate, and closely reasoned account of current research, making it more accessible to a wider audience. Further research could include similar treatment of Klima#s tunnels, which have provided the fastest known attack on MD5.

Summary Statement
This paper expands on one of the most comprehensive papers on the differential attack on MD5 by providing original proofs, corrections, and illustrations to make the attack more accessible to the mathematically-literate reader.

Help Received
Qualcomm employee helped with concepts; father helped edit the report.