The Complexity of Mendel and Thomsen's Preimage Attack on JH-512

Hongjun Wu

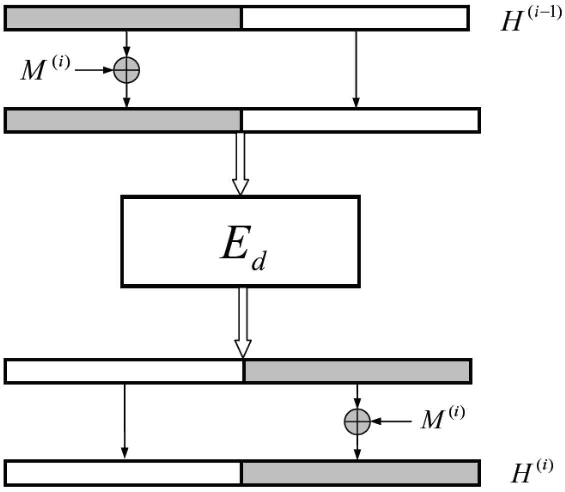
Institute for Infocomm Research, Singapore
wuhongjun@gmail.com

Abstract

Mendel and Thomsen gave a preimage attack on JH-512 by finding a preimage through the collision search over the space of 2^{1024} elements. However, they did not estimate the cost of the collision search which is the most expensive part in their attack. Our analysis shows that their attack requires at least $2^{510.3}$ compression function computations, $2^{510.6}$ memory ($2^{516.6}$ bytes), 2^{524} memory accesses and 2^{524} comparisons. Such complexity is far more expensive than brute force attack which requires 2^{512} compression function computations and almost no memory.

1 The JH Compression Function [2]

The JH structure is extremely simple. F_8 is the compression function used in hash function JH. F_8 compresses the 512-bit message block $M^{(i)}$ and 1024-bit $H^{(i-1)}$ into the 1024-bit $H^{(i)}$. F_8 is constructed from the bijective function E_8 as shown below:



2 Mendel and Thomsen's Preimage Attack on JH-512 [1]

Mendel and Thomsen's Preimage attack works on messages with 4 message blocks. It is shown below:

1. Choose an arbitrary message block M_4 with correct padding, and compute $H_3 = f^{-1}(x \| h, M_4$ for an arbitrary 512-bit value x.
2. Compute 2^{509} candidates for $H_2 = f^{-1}(H_3, M_3)$ with arbitrary choices of M_3, and save the pairs (H_2, M_3) in a list L.
3. Use M_1 to construct an r-collision for the 512 higher bits of H_1, given the initial value H_0 of JH-512. For $r = 51$ this has a complexity of about $2^{506.3}$ compression function evaluations. In other words, we find $r = 51$ message blocks M_1^k for $0 \leq k < r$ such that b^k is equal with $H_1^k = a^k \| b^k$.
4. Compute $\Delta^k = H_1^0 \oplus H_1^k$ for $0 \leq k < r$.
5. Choose an arbitrary message block M_2 and compute $H_2 = f(H_1^0, M_2)$ and check if $H_2^k = H_2 \oplus \Delta^k$ for $0 \leq k < r$ is in the list L. The probability for each choice of M_2 is about $51 \cdot 2^{1024-509}$, so we need to try an expected $2^{515}/51 \approx 2^{509.3}$ message blocks. Note that only about $2^{512}/51 \approx 2^{506.3}$ different message blocks can be chosen in this step without repetition, and hence we must find an expected 2^3 51-collisions in step 3. However, 2^3 51-collisions can be found in time only a factor about $2^{3/51} \approx 2^{0.06}$ more than a single 51-collision. Thus, the "new" complexity of step 3 is $2^{506.3}$ (unchanged to one decimal place), and the current step has complexity about $2^{509.3}$ (we ignore the 51 xors needed in this step, assuming this takes negligible time compared to one evaluation of f).
6. Once we have found H_2^k such that a pair (H_2^k, M_3) is in the list L, we have to adjust M_1 and M_2 accordingly such that $f(f(H_0, M_1), M_2) = H_2^k = H_2 \oplus \Delta^k$.
   It is easy to see that this can be achieved by setting $M_1 = M_1^k$ and $M_2 = M_2 \oplus \Delta^k$, since:

$$H_1 = f(H_0, M_1^k) = H_1^k = H_1^0 \oplus \Delta^k$$
$$H_2 = f(H_1^0 \oplus \Delta^k, M_2 \oplus \Delta^k) = H_2 \oplus \Delta^k = H_2^k$$

Note that there is an error in the Step 5 above. For $2^{506.3}$ choices of M_2, the number of different H_2 being generated would be less than 2^{512}. The reason is that by choosing $2^{506.3}$ different M_2, it is impossible for $M_2 \oplus \Delta^k$ $(0 \leq k < r)$ to generate 2^{512} different values since the first 512 bits of Δ^k are random. To fix this error, the attacker has to generate more sets of r-collision. Thus more computation is needed in the attack.

3 The Complexity of Mendel and Thomsen's Preimage Attack on JH-512

Before analyzing the complexity of Mendel and Thomsen's preimage attack on JH-512, we give the complexity of sorting and the complexity of checking whether an element is in an sorted table.

The complexity of sorting. There are a number of sorting algorithms. The complexity of sorting n random elements is at least $O(n{\cdot}\log n)$. Different algorithms requires different amount of memory accesses and comparisons.

The complexity of checking whether an element is in an sorted table. It takes about log$n$ memory accesses and comparisons to determine whether an element is in an sorted table with n elements.

We give below the detailed complexity analysis of Mendel and Thomsen's preimage attack on JH-512.

1. Step 1 requires negligible amount of computation.
2. Step 2 requires 2^{509} compression function computations and $2^{509} \times 3 = 2^{510.6}$ memory (each unit of memory is 64 bytes).
3. Step 3 requires $2^{506.3}$ compression function computations, $2^{506.3} \times 3 = 2^{507.9}$ memory, and the sorting of a table with $2^{506.3}$ elements. The reason is that to find 51-collision for the 512 higher bits of H_1, it requires the sorting of $2^{506.3}$ elements.
4. Step 4 requires negligible amount of computation.
5. Step 5 is the most expensive part of the attack. Firstly, the attacker must sort the table L that is generated in Step 2. Note that table L consists of 2^{509} elements. Secondly, the attacker must perform at least $2^{506.3} \times 2^3 = 2^{509.3}$ compression function computations in order to generate 2^{515} different H_2. Thirdly, the attacker must find out which H_2 is in the table L. Checking whether one H_2 is in L or not requires about 2^9 memory accesses and comparisons. It requires $2^{515} \times 2^9 = 2^{524}$ memory accesses and comparisons to find the collision.
6. Step 6 requires negligible amount of computation.

Thus the complexity of Mendel and Thomsen's preimage attack on JH-512 is:

1. compression function computations: $2^{509} + 2^{506.3} + 2^{509.3} = 2^{510.3}$.
2. memory: $2^{510.6}$ ($2^{516.6}$ bytes) (The memory in Step 3 can be ignored when it is performed before Step 2.)
3. sorting: the attack requires the sorting of two tables, one table with 2^{506} elements, another table with 2^{509} elements.
4. memory accesses (excluding those required in sorting): 2^{524}
5. comparisons (excluding those required in sorting): 2^{524}

Considering that there is an error in Mendel and Thomsen's preimage attack (as pointed out in Section 2), the complexity would be even higher.

Let us take a look at the brute force preimage attack on JH-512. The attack requires 2^{511} hash function computations to find a preimage in average. Each hash function computation requires at least two compression function computations (due to the padding of JH). Thus finding a preimage of JH-512 requires about 2^{512} compression function computations in average. Almost no memory is requires in the brute force attack.

4 Conclusion

The complexity of Mendel and Thomsen's attack is much higher than that of brute force attack. Although the number of compression function computations is reduced to $2^{510.3}$, the number of memory accesses and comparisons increases to 2^{524}, and the memory increases to $2^{516.6}$ bytes. The time complexity of the memory accesses and comparisons in their attack already exceeds that of brute force attack. The memory cost makes their attack far more worse than brute force attack.

The complexity analysis of Mendel and Thomsen's preimage attack on JH-512 shows that the collision search over the space of 2^{1024} elements is the most expensive part in their attack. The 1024-bit hash value in JH plays an important role in defending such attack. It shows that the extremely simple structure of JH is strong against this type of dedicated attack.

References

[1] F. Mendel and S. S. Thomsen, "An Observation on JH-512". Available at http://ehash.iaik.tugraz.at/uploads/d/da/Jh_preimage.pdf .

[2] H. Wu. The Hash Function JH. Submission to NIST, 2008. Available at http://icsd.i2r.a-star.edu.sg/staff/hongjun/jh/jh.pdf.