Interactions Between Data Compression and Encryption

James Iwamasa

February 12, 2017

Abstract

Every day, insurmountable amounts of data are shuffled around over an expansive web of users. While the infrastructure to facilitate such movements of information is nothing short of a world wonder, this system inherently poses many problems, the two biggest being that of bandwidth and security. While the fields of data compression and encryption are well developed and highly specialized in their own rights, they both involve the act of altering the raw data somehow, and as a result can come into conflict. This paper will explore these complex interactions between data compression and encryption.

1 Introduction

The direction of this paper can be characterized by a single question: Do we compress our data before we encrypt or the other way around? Basic knowledge of standard encryption and compression techniques will imply that, typically, compression should come first. Data compression often works by noticing patterns and redundancies in our data. On the other hand, encryption attempts to make our data unrecognizable gibberish to anyone who doesn't know the secret code. Thus, if we want our compression to have the most effect, we should not encrypt it first.

One may also have the intuition that compression might actually assist in encryption, since we're still converting our data into something only a computer can effectively decompress. But there have been studies[1][2] showing that compression can actually expose security flaws. By examining the changes in the length of the data from the compression, hackers can actually infer the plaintext of http requests (as in expoits CRIME and BREACH).

This conflict of interests between encryption and compression is the main crux of this paper. Now we will explore how the two can interfere with each other, peacefully coexist, and perhaps even help each other in protecting and packaging our data.

2 Possible Topics to Cover

Overview of encryption/methods.

Overview of data compression/methods.

Conflicts (CRIME, BREACH, etc)

How they can work well with each other

References

[1] J. Kelsey. Compression and Information Leakage of Plaintext. In *Fast Software Encryption, 9th International Workshop, FSE 2002, Leuven, Belgium, February 4-6, 2002, Revised Papers*, volume 2365 of *Lecture Notes in Computer Science*, pages 263-276. Springer, February 2002.

[2] Y. Gluck, N. Harris, and A. Prado. BREACH: Reviving the CRIME Attack. URL: http://breachattack.com/resources/BREACH. [cited February 2017].

[3] R. Sharma and S. Bollavarapu. Data Security using Compression and Cryptography Techniques. In *International Journal of Computer Applications (0975-8887)*, volume 117 - No. 14, May 2015.

[4] M. Johnson, P. Ishwar, V. Prabhakaran, D. Schonberg, and K. Ramchandran. On Compressing Encrypted Data. In *IEEE Transactions on Signal Processing*, volume 53 - No. 10, October 2004.