Cryptology ePrint Archive: Report 2011/088

Leftover Hash Lemma, Revisited

Boaz Barak and Yevgeniy Dodis and Hugo Krawczyk and Olivier Pereira and Krzysztof Pietrzak and Francois-Xavier Standaert and Yu Yu

Abstract: The famous Leftover Hash Lemma (LHL) states that (almost) universal hash functions are good randomness extractors. Despite its numerous applications, LHL-based extractors suffer from the following two drawbacks:

(1) Large Entropy Loss: to extract v bits from distribution X of min-entropy m which are e-close to uniform, one must set v <= m - 2*log(1/e), meaning that the entropy loss L = m-v >= 2*log(1/e).

(2) Large Seed Length: the seed length n of (almost) universal hash function required by the LHL must be at least n >= min(u-v, v + 2*log(1/e))-O(1), where u is the length of the source.

Quite surprisingly, we show that both limitations of the LHL --- large entropy loss and large seed --- can often be overcome (or, at least, mitigated) in various quite general scenarios. First, we show that entropy loss could be reduced to L=log(1/e) for the setting of deriving secret keys for a wide range of cryptographic applications. Specifically, the security of these schemes gracefully degrades from e to at most e + sqrt(e * 2^{-L}). (Notice that, unlike standard LHL, this bound is meaningful even for negative entropy loss, when we extract more bits than the the min-entropy we have!) Based on these results we build a general *computational extractor* that enjoys low entropy loss and can be used to instantiate a generic key derivation function for *any* cryptographic application.

Second, we study the soundness of the natural *expand-then-extract* approach, where one uses a pseudorandom generator (PRG) to expand a short "input seed" S into a longer "output seed" S', and then use the resulting S' as the seed required by the LHL (or, more generally, any randomness extractor). Unfortunately, we show that, in general, expand-then-extract approach is not sound if the Decisional Diffie-Hellman assumption is true. Despite that, we show that it is sound either: (1) when extracting a "small" (logarithmic in the security of the PRG) number of bits; or (2) in *minicrypt*. Implication (2) suggests that the sample-then-extract approach is likely secure when used with "practical" PRGs, despite lacking a reductionist proof of security!

Finally, we combine our main results to give a very *simple and efficient* AES-based extractor, which easily supports variable-length messages, and is likely to offer our *improved entropy loss bounds* for any computationally-secure application, despite having a *fixed-length* seed.

Category / Keywords: foundations / Leftover Hash Lemma, Randomness Extractors, Key Derivation, Pseudorandom Generators, Entropy Loss.

Date: received 26 Feb 2011, last revised 3 Sep 2011

Contact author: dodis at cs nyu edu

Available formats: Postscript (PS) | Compressed Postscript (PS.GZ) | PDF | BibTeX Citation

Version: 20110903:194557 (All versions of this report)

Discussion forum: Show discussion | Start new discussion

[Cryptology ePrint archive]