



Research Article

Cryptographic Competitions*

Daniel J. Bernstein

Department of Computer Science, University of Illinois at Chicago, Chicago, IL60607–7045, USA
Horst Görtz Institute for IT Security, Ruhr University Bochum, Bochum, Germany
djb@cr.yp.to

Communicated by Tetsu Iwata

Received 21 September 2020 / Revised 24 April 2023 / Accepted 30 April 2023

Abstract. Competitions are widely viewed as the safest way to select cryptographic algorithms. This paper surveys procedures that have been used in cryptographic competitions, and analyzes the extent to which those procedures reduce security risks.

Keywords. Cryptography, Competitions, DES, AES, eSTREAM, SHA-3, CAESAR, NISTPQC, NISTLWC.

1. Introduction

The CoV individual reports point out several shortcomings and procedural weaknesses that led to the inclusion of the Dual EC DRBG algorithm in SP 800-90 and propose several steps to remedy them. ...

The VCAT strongly encourages standard development through open competitions, where appropriate. —“NIST Cryptographic Standards and Guidelines Development Process: Report and Recommendations of the Visiting Committee on Advanced Technology of the National Institute of Standards and Technology” [133], 2014

Cryptographic competitions are not a panacea. DES, the output of the first cryptographic competition, had an exploitable key size (see [57], [73], [140], [36], and [63]), had an exploitable block size (see [96] and [35]), and at the same time had enough denials

*This work was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany’s Excellence Strategy—EXC 2092 CASA—390781972 “Cyber Security in the Age of Large-Scale Adversaries,” by the US National Science Foundation under grant 1913167, and by the Cisco University Research Program. “Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation” (or other funding agencies). Permanent ID of this document: b7af715576cc229aaf8c532ea89bb6ace1c91a65. Date: 2023.04.24.

of exploitability (see, e.g., [74], [55, Section 7], [76], and [1]) to delay the deployment of stronger ciphers for decades. As another example, AES performance on many platforms relies on table lookups with secret indices (“S-table” or “T-table” lookups), and these table lookups were claimed to be “not vulnerable to timing attacks” (see [54, Section 3.3] and [102, Section 3.6.2]), but this claim was incorrect (see [17] and [129]), and this failure continues to cause security problems today (see, e.g., [47]). As a third example, SHA-3 was forced to aim for a useless 2^{512} level of preimage security, and, as a result, is considerably larger and slower than necessary, producing performance complaints and slowing down deployment (see, e.g., [89])—which is a security failure if it means that applications instead use something weak (see, e.g., [92]) or nothing at all.

I don’t mean to suggest that competitions are a bad idea. I can think of many more failures of cryptographic algorithms selected *without* competitions. But it is surprisingly difficult to find literature systematically analyzing the security risks in various algorithm-selection processes, and systematically working on designing processes that reduce these risks. Even if competitions are the best approach, there are many different types of competitions, and we should understand how these variations can avoid—or create—security risks.

This paper surveys, from the perspective of a skeptical security reviewer, the procedures used in cryptographic competitions. Given my role in organizing CAESAR, I have used CAESAR as a running example in the paper—among other things, reporting how CAESAR started, how it was run, and what it produced—but I have also used other cryptographic competitions as examples, including the DES, AES, eSTREAM, SHA-3, and NISTLWC competitions for symmetric algorithms, and the NISTPQC competition for asymmetric (public-key) algorithms. I hope that the analysis here is of value for future competitions.

1.1. *Related Work*

For per-competition reports (at various levels of detail and finality) on other competitions, see [55] for DES; [103] and [102] for AES; [113] for eSTREAM; [111, 132], and [45] for SHA-3; [130] and [131] for NISTLWC; and [8, 9], and [10] for NISTPQC. Beyond these reports, there is a vast literature on the security and performance of various submissions to the competitions.

The *procedures* used in cryptographic competitions are frequently mentioned as introductory material, but not as a core topic of cryptographic risk analysis. NIST’s description of its cryptographic standardization process [53] includes a few paragraphs summarizing competition processes, saying that “competitions can focus the attention of cryptographers around the world.” A deeper discussion of security-evaluation and performance-evaluation procedures—providing many reasons to wonder how effective competitions really are at reducing security risks—had appeared in [109, Sections 2 and 4] as part of input to the AES process from NESSIE, a joint project by European cryptography experts.

Security failures in cryptographic algorithms are traditionally blamed upon the specific algorithms that failed, and not upon the processes that selected those algorithms: cryptographers recommend against the algorithms while continuing to use the same processes. However, there has been some attention to the idea of protecting standardization

processes against sabotage: see [133], [28], [34], [61], and [26]. A standardization process can fail even when it is not under attack; in [23, Appendix B] I called for a systematic study of how reliably a standardization process produces secure standards. The idea that different processes can create different levels of risks was already a prerequisite for the common belief that competitions are less risky than typical standardization processes. The same idea is standard in the broader literature on risk analysis: see, e.g., [106].

2. Speed

A competition is defined by the Collins English Dictionary as “an event in which many people take part in order to find out who is best at a particular activity.”

One of the traditional forms of competition is a speed competition—a race. Who can swim the fastest? Who can run the fastest? Who can write a sorting program that runs the fastest? Who can drive the fastest? Who can build the fastest car? Who can write the fastest encryption program? Note that one can suppress the role of the humans in some of these questions, thinking of the cars as the competitors in a car race and thinking of the programs as the competitors in a software competition.

Sometimes speed competitions have a clear utilitarian purpose. Ancient Greek legend says that a runner ran 40 kms to Athens to report success in the Battle of Marathon. In selecting a runner for such a task, the commander wants to know, first, who can complete the task at all, and, second, who can complete it in the required amount of time. It would not be surprising if the original Marathon runner was selected as someone who did well enough in a previous race. Note that managers who have no idea what the required amount of time is will want to take specifically the *winner* of the competition, so as to reduce the risk of being too slow; someone who knows more about the requirements will tend to think that taking the winner is less important.

Sometimes people participate in or watch speed competitions simply for the thrill. People participating in a marathon today can reasonably claim health benefits, but people *watching* a marathon on the living-room television don’t have this excuse. Nobody claims that we need to know the fastest runner so as to decide who should be carrying a message to Athens.

If I set a speed record for some computation, am I doing it just for the thrill? Does the speed record actually matter for users? Making software run faster is a large part of my research; I want to *believe* that this is important, and I have an incentive to exaggerate its importance. People who select my software for its performance have a similar incentive to exaggerate the importance of this performance as justification for their decisions. Perhaps there is clear evidence that the performance is important, or perhaps not.

2.1. *The Machinery of Cryptographic Performance Advertising, Part 1: Measurements*

In 2019, a Google blog post [52] (see also the accompanying paper [51]) presented the following convincing story:

- Android had required storage encryption since 2015 *except* on “devices with poor AES performance (50 MiB/s and below).”

- For example, on an ARM Cortex-A7, storage encryption with AES is “so slow that it would result in a poor user experience; apps would take much longer to launch, and the device would generally feel much slower.” (According to [13], the Cortex-A7 “has powered more than a billion smartphones.”)
- Starting in 2019, Android Pie was enabling storage encryption for these devices using Adiantum, a wide-block cipher built on top of ChaCha12.
- On a 1.19GHz Cortex-A7, AES-256-XTS decrypts at 20 MB/s. Adiantum decrypts at 112 MB/s.

What follows is another example of one cryptographic system solving the performance problems that caused another cryptographic system to be rejected.

Your Internet service provider used the Domain Name System (DNS) in three steps to find the address of `www.google.com`. It learned the address of `www.google.com` from the `google.com` servers; earlier it learned the address of the `google.com` servers from the `.com` servers; earlier it learned the address of the `.com` servers from the Internet’s central “root DNS servers.” Each address is cached for some time, reducing the overall load on the root DNS servers to 120 billion queries per day (according to [120]), around 1.4 million queries per second. These servers have 13 names (root server A, root server B, and so on through root server M), but according to [116] there are actually “hundreds” of physical computers; if “hundreds” means 200 then an average root server handles around 7000 queries per second.

A 2021 statement [117] from the root operators indicates that these servers will not deploy encryption for now. The primary reason stated is performance:

Due to the critical role that root name servers play, combined with the fact that they are themselves often targets of DDoS attacks, Root Server Operators have some concerns about supporting DNS encryption for serving the root zone. It is well known that UDP has desirable performance characteristics, due to its stateless nature. Increasing the state-holding burden with the addition of connection-oriented protocols, as well as encryption data, not only reduces the performance of name servers, but also may raise new types of denial-of-service attacks.

It is certainly true that running DNS over a stateful HTTPS connection, rather than over stateless UDP, opens up denial-of-service attacks; see, e.g., [33, Section 2.2]. But it is simply not true that DNS encryption needs connections. DNSCurve, a simple DNS transport layer that uses X25519 for encryption and authentication, was already deployed by OpenDNS in 2010, according to [104]; there are faster DH proposals, but X25519 suffices for the following analysis. DNSCurve runs over UDP in the same way that DNS normally does, rather than requiring the server to accept connections and store state for each connection.

Regarding the impact of encryption on CPU load, the software from [98] takes 95437 Skylake cycles for X25519, and all other operations in DNSCurve are much faster. The software performs more than 125000 X25519 operations per second on a low-cost quad-core 3GHz Intel Xeon E3-1220 v5 CPU from 2015. If a root server is using such a CPU

to handle its 7000 DNS queries per second, and those queries all upgrade tomorrow to DNSCurve, then X25519 will consume under 6% of the available CPU time. Presumably most root CPUs are more powerful than this; on a newer, higher-cost server with two AMD EPYC 7742 CPUs from 2019, the same software performs more than 2.9 million operations per second.

Regarding network traffic, a DNS query for `.com` today is 21 bytes plus 8 bytes of UDP overhead, 20 bytes of IP overhead, and 38 bytes of Ethernet overhead; the response is 509 bytes (without IPv6) plus the same overheads. DNSCurve makes each query 68 bytes longer—there is an 8-byte protocol selector, a 32-byte X25519 key, a 12-byte nonce, and a 16-byte MAC—and each response 48 bytes longer. This 17.5% difference in traffic volume is barely noticeable compared to the much larger network capacity already available in the root servers. To quantify this, consider the 2015 denial-of-service attack described in [115], an attack that “saturated network connections” for *some* of the root servers but still did not take down root DNS service. This attack used “up to approximately 5 million queries per second, per DNS root name server letter receiving the traffic,” and hit “most” of the 13 letters—tens of millions of queries per second overall. Presumably the root servers have even more bandwidth today.

What happens if, instead of trying to deny service by flooding a network, an attacker tries to deny service by flooding a CPU with X25519 operations? The software mentioned above handles 750 megabits per second of X25519 keys on the EPYC server. This corresponds to more than 3 gigabits per second of DNS network traffic, given minimum packet overheads. Sites with larger Internet connections can spread the load across multiple servers to guarantee that the network running at full capacity cannot overload the CPUs with cryptographic operations. Sites with smaller Internet connections can use smaller servers.

2.2. *The Machinery of Cryptographic Performance Advertising, Part 2: Confirmation Bias*

Usually the factual basis for cryptographic performance advertising is much less clear. Here’s a case study.

The 2020 paper “Post-quantum authentication in TLS 1.3: a performance study” [124] states “Reports like [1] lead us to believe that hundreds of extra milliseconds per handshake are not acceptable.” The cited document “[1]” is a 2017 Akamai press release “Akamai online retail performance report: milliseconds are critical” [6], subtitled “Web performance analytics show even 100-millisecond delays can impact customer engagement and online revenue.”

Akamai’s underlying report [7] says, as one of its “key insights,” that “just a 100-millisecond delay in load time hurt conversion rates by up to 7%.” My understanding is that “conversion” has the following meanings:

- in traditional sales terminology, converting a potential customer into a lead (i.e., contact information for the potential customer) or a sale;
- in web sales terminology, converting a view of a product web page into a sale of the product.

A reader who digs into the report finds a statement that “desktop pages” that loaded “100 milliseconds slower” experienced a “2.4% decrease in conversion rate”; and a statement that for smartphones 2.4% changes to 7.1%. Apparently these statements are the source of the “key insight” stating “up to 7%.”

Let’s look at the underlying data more closely. Akamai hosts its customers’ web pages, caching copies of those pages on thousands of Akamai-run computers around the world, the idea being that browsers will receive each page quickly from a nearby computer. The report says that Akamai collected metadata on billions of web sessions from “leading retail sites” that are Akamai customers (with permission from the sites). In Fig. 1 here, a screenshot from [7, page 8], the yellow area shows the distribution of page-load times in converted web sessions, the blue area shows the same for non-converted web sessions, and the red curves show the percentage of web sessions that were converted.

For example, for desktop browsers (top graph), about 12% of sessions were converted for page-load times around 1.8 s: at horizontal position 1.8, the red curve is around 12%, and the top of the yellow is about 12% of the top of the blue. The conversion rate drops off for larger page-load times: e.g., about 6% of sessions were converted for page-load times around 4 s, meaning that the conversion rate was 50% smaller. The report highlights the conversion rate being 12.8% at 1.8 s, and 2.4% smaller (i.e., 12.5%) at 1.9 s, the conclusion being that a 100-millisecond delay in load time is measurably bad.

Notice, however, that the conversion rate in the graph also drops off for *smaller* page-load times. Compare the following numbers:

- ≈6% of sessions were converted for page-load times around 1.1 s.
- ≈8% of sessions were converted for page-load times around 1.2 s.
- ≈9% of sessions were converted for page-load times around 1.3 s.
- ≈12% of sessions were converted for page-load times around 1.8 s.

Why did Akamai’s report not conclude that a 100-millisecond delay in load time is measurably *good*?

The report briefly addresses this question, saying “the faster end of the bell curve is primarily comprised of 404 pages and other pages that, while fast, do not fall on the conversion path.” But wait a minute. If different types of web pages can fail to produce sales, explaining the low end of the graph, then can’t such differences explain the *entire* graph? Maybe the situation is that there are

- “too simple” product pages (which tend to load quickly) that don’t have enough pictures to convince the typical customer, and
- “too complex” product pages (which tend to load slowly) that scare the typical customer away, and
- “just right” product pages (which tend to have load times in the middle) that do the best job of attracting customers.

Could it be that what matters for sales isn’t actually the last bit of speed in delivering the web page, but rather the content of the web page?

Another section of the same report says that “user sessions that converted contained 48% more scripts than sessions that did not convert.” Perhaps a three-dimensional analysis of conversion, scripts, and load times would show that load times don’t matter when the number of scripts is the same. Perhaps there are other confounding factors, such as lower-income customers buying fewer products and also tending to have slower network connections.

Recall that Akamai’s report portrays a 100-millisecond delay as being worse for smartphones than for desktops, losing 7.1% of the smartphone conversions. But the smartphone conversion rate in Fig. 1 (middle red graph) drops off *more gently* than the desktop conversion rate. For example, the smartphone conversion rate is 3.3% at 2.7 s, and is half of 3.3% around 6 s. Akamai’s report seems to be alluding to the first drop in the red graph after its peak, but the red graph then jumps up a moment later, and in general the small-scale wobbling in the red graph looks like random noise.

Maybe a properly designed study that adds 100 milliseconds of delay into web pages for randomly selected users would produce a 2%, maybe even 7%, drop in sales. But the study in [7] was not properly designed. There are many other theories that would produce the graphs in [7]. The highlighting of one *possible* theory is easily explained as a combination of confirmation bias and marketing.¹ The same report briefly mentions that “Walmart saw up to a 2% increase in conversions for every second of improvement in load time,” without noting that, compared to 7% for 100 milliseconds, 2% for 1 s is $35\times$ less important.

2.3. A Screenshot

See Fig. 1.

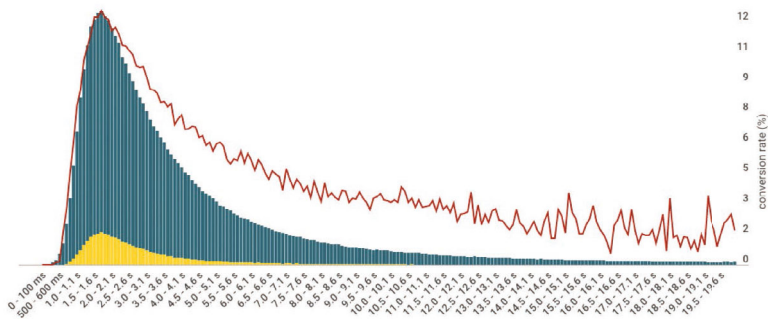
2.4. The Machinery of Cryptographic Performance Advertising, Part 3: Systematic Exaggeration

Let’s return to the belief in [124] that “hundreds of extra milliseconds per handshake are not acceptable.” Most of the page loads in Fig. 1 are at least hundreds of milliseconds beyond the red peaks, and yet these are, according to Akamai, deployed web pages from “leading retail sites”; so how does [124] conclude that this extra time is “not acceptable”?

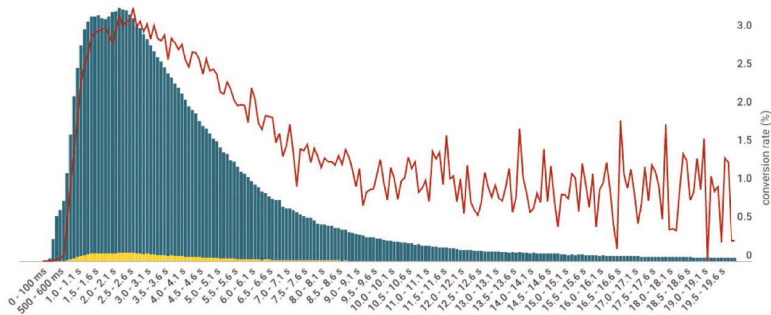
Even if hundreds of extra milliseconds per page load are unacceptable, it is an error to conflate this with hundreds of extra milliseconds *per handshake*. A TLS handshake sets up a session that can be, and often is, used to load many pages without further handshakes. A page often collects content from multiple servers, but presumably most latencies overlap. Perhaps users wouldn’t actually notice the costs considered in [124]—or perhaps they would notice the costs but would consider the costs *acceptable* if sufficient benefit is provided.

¹I’m not trying to say that Akamai’s many customers are making a mistake. On the contrary: my impression is that Akamai is providing robust web service to its customers, and at the same time is doing a valuable public service in allocating Internet links more efficiently.

How does load time correlate to conversion rate? (desktop)



How does load time correlate to conversion rate? (mobile)



How does load time correlate to conversion rate? (tablet)

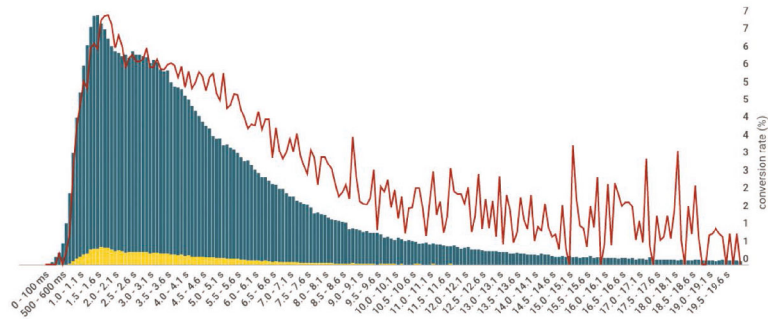


Fig. 1. Screenshot from [7, page 8].

Google constantly makes changes to its Chrome browser. Sometimes these changes improve performance. Sometimes they reduce performance. Google has a statement [67] of procedures for deciding whether a performance regression is acceptable. This statement spends several paragraphs listing “some common justification scenarios,” such as the following:

What do we gain? It could be something like: ... Additional security

Compare this to a November 2021 Cloudflare blog post [139] pointing to the same statement of Google procedures and claiming that “only in exceptional cases does Chrome allow a change that slows down any microbenchmark by even a percent.” This claim plays a pivotal role in the advertising for [139], which, like [124], studies the performance of signature systems.

All of the signature systems listed in [124, Table 1] have software available that signs in under 20 milliseconds on a 3GHz Intel Haswell CPU core from 2013 (never mind the possibility of parallelizing the computation across several cores). The server signs a TLS handshake only once. The browser also verifies certificate chains—the paper considers TLS sessions with 3 or 4 verification operations—but all of these signature systems have software available that verifies in under 5 milliseconds. The total size of a public key and a signature in [124, Table 1] is at most 58208 bytes, so a 100Mbps network connection (already common today, never mind future trends) can transmit 4 public keys and 4 signatures in 20 milliseconds. For comparison, [75, “Total Kilobytes”] shows that the median web page has grown beyond 2MB and that the average is even larger.

Given the numbers, it is hard to see how TLS users will care which of these signature systems is used, and it is hard to see why one should care about a more detailed performance study. The paper [124] thus has an incentive to paint a different picture. The paper starts from the “not acceptable” belief quoted above; selects signature-system software that was “not optimized”; configures a server-controlled networking parameter, `initcwnd`, to wait for a client reply after sending just 15KB of data; and devotes a page to discussing long-distance connections, such as a US-to-Singapore connection, where waiting for a client reply costs a 225-millisecond round trip.²

The server’s `initcwnd` sheds light on the question of how important latency is. In the original TCP congestion-control algorithms from the late 1980s [79], a server with any amount of data to send begins by sending at most one packet to the client. The server then waits for a client reply, then sends a burst of two packets, then waits for a client reply, then sends a burst of four packets, etc. The initial packet is allowed to be full size, which today typically means 1500 bytes. The choice to start with one full-size packet, not more and not less, is a balance between (1) the slowdown from delaying subsequent data and (2) concerns that having everyone start with more data would overload the network.

In 1998, an “Experimental” RFC [11] proposed increasing the server’s “initial congestion window” (`initcwnd`) from 1 packet to 2–4 packets. Today this is a “Proposed Standard” [12]. In 2013, another “Experimental” RFC [46] proposed increasing `initcwnd` to 10 packets. According to [44], Akamai was using 16 packets in 2014, and 32 packets in 2017. Five doublings of `initcwnd`, from 1 to 2 to 4 to 8 to 16 to 32, eliminate four or five round-trip times from any sufficiently large file transfer.

Other major web servers in 2017 used `initcwnd` ranging from 10 through 46, according to [44]. All of these are still officially “experimental,” far above the standard 1 and the proposed standard 2–4, but most Internet links have grown to handle massive video traffic, and a relatively tiny burst of packets at the beginning of a TCP connection

²The blog post [139] instead measures Cloudflare’s usual short-distance connections. Unsurprisingly, the observed times are much smaller than in [124].

does not cause problems. Meanwhile [124] refuses to benchmark `initcwnd` above 10, and issues an ominous warning that widespread deployment of a larger `initcwnd` “could have adverse effects on TCP Congestion Control,” as if a larger `initcwnd` were not already widely deployed.

In the opposite direction, compared to a web server taking `initcwnd` as 46, a web server taking `initcwnd` as just 10 is sacrificing two round trips, almost half a second for a connection between the USA and Singapore. If such a slowdown is “not acceptable” then why are some major web servers doing it? Perhaps the answer is that such long-distance connections are rare. Or perhaps the answer is that occasional delays of hundreds of milliseconds aren’t actually so important.

2.5. Competitions for Cryptographic Performance

There is overwhelming evidence of performance requirements—whether real or imagined—playing an important, perhaps dominant, role in cryptographic competitions:

- At an NBS workshop in 1976, before DES was approved as a standard, Diffie (in joint work with Hellman; see [57, page 77, “Cost of larger key”]) proposed modifying the DES key schedule to use a longer key. Representatives of Collins Radio and Motorola objected to this proposal, saying that DES is “close to the maximum that could be implemented on a chip with present technology” and that a manufacturing delay “of one to two years might be encountered if a longer key were required.” See [97, page 20].³
- The AES call for submissions [87] listed “computational efficiency” as the second evaluation factor after “security.” NIST’s final AES report [102, page 528] stated that “Rijndael appears to offer an adequate security margin” and that “Serpent appears to offer a high security margin,” and the same report claimed [102, page 516] that “security was the most important factor in the evaluation,” but NIST selected Rijndael rather than Serpent as AES. NIST’s only complaints about Serpent were performance complaints.
- eSTREAM called [62] for “stream ciphers for software applications with high throughput requirements” and called for “stream ciphers for hardware applications with restricted resources such as limited storage, gate count, or power consumption.” The eSTREAM committee selected several ciphers for the final eSTREAM portfolio [15]—for example, selecting my Salsa20 cipher with 12 rounds, “combining a very nice performance profile with what appears to be a comfortable margin for security.” I had recommended, and continue to recommend, 20 rounds; I had proposed reduced-round options only “to save time” for users “who value speed more highly than confidence.”

³NBS wrote in 1977 [55, page 10] that DES was “satisfactory for the next ten to fifteen years as a cryptographic standard.” NIST did not end up withdrawing DES as a standard until 2005. Almost all of the DES benefits claimed in [90] appeared in 1980 or later, as did 93% of the implementations listed in [90, page 31]. Why was a claimed manufacturing delay from 1976 to 1977 or 1978 treated as important? See Sect. 3.6 for a possible answer. It is also far from clear that the claim was correct: [90, page 16] reports that IBM “had developed a commercially viable VLSI chip that could incorporate the encryption algorithm efficiently” already before March 1975.

- The SHA-3 call for submissions [84] said that “NIST expects SHA-3 to have a security strength that is at least as good as the hash algorithms currently specified in FIPS 180–2, and that this security strength will be achieved with significantly improved efficiency.” When NIST selected Keccak as SHA-3, it wrote [45] that Keccak “offers exceptional performance in areas where SHA-2 does not,” and that Keccak “received a significant amount of cryptanalysis, although not quite the depth of analysis applied to BLAKE, Grøstl, or Skein.” On the other hand, NIST’s prioritization of efficiency over security was not as clear for SHA-3 as for AES: [45] also said that Keccak “relies on completely different architectural principles from those of SHA-2 for its security.”
- CAESAR called [19] for “authenticated ciphers that (1) offer advantages over AES-GCM and (2) are suitable for widespread adoption.” Proposals varied in whether they emphasized security advantages or efficiency advantages. The CAESAR committee later identified three “use cases” [21]: “lightweight applications” requiring “small hardware area and/or small code for 8-bit CPUs”; “high-performance applications” requiring “efficiency on 64-bit CPUs (servers) and/or dedicated hardware”; and, deviating from the speed theme, “defense in depth” providing “authenticity despite nonce misuse.” Ultimately, the CAESAR committee selected Ascon (first choice) and ACORN (second choice) for use case 1, AEGIS-128 and OCB (without an order) for use case 2, and Deoxys-II (first choice) and COLM (second choice) for use case 3.
- The NISTPQC reports referred repeatedly to performance as a basis for decisions. For example, [9] said that NIST’s “first priority for standardization is a KEM that would have acceptable performance in widely used applications overall. As such, possible standardization for FrodoKEM can likely wait.” What is not “acceptable” in Frodo’s performance? NIST wrote that a TLS server using Frodo uses “close to 2 million cycles” and “receives a public key and a ciphertext (around 20,000 bytes in total) for every fresh key exchange.”⁴ Similarly, the report [10] on the first batch of NISTPQC winners (Dilithium, Falcon, Kyber, and SPHINCS+) praised Frodo’s “conservative design choices” but claimed that Frodo “is clearly not an immediate drop-in general-purpose scheme.”
- NISTLWC called [100] for submissions of hash functions and authenticated ciphers “tailored for resource-constrained devices,” since “many conventional cryptographic standards” are “difficult or impossible to implement” in such devices, at least with the constraint of acceptable performance. NISTLWC is the most recent competition in this list, with an initial submission deadline in 2019; NISTLWC selected Ascon in 2023.

Almost all of these competitions are for symmetric cryptography: block ciphers, hash functions, authenticated ciphers, etc. Symmetric cryptography is applied to *every byte of communicated data*, authenticating every byte and encrypting every potentially confidential byte. The Android storage encryption example (almost) fits this pattern, with

⁴One side would send a 9616-byte public key, and the other side would send back a 9720-byte ciphertext, so “receive” is not true. It’s true that each side would use close to 2 million Haswell cycles: i.e., close to a full *millisecond* on a 2GHz CPU core.

every byte encrypted (but currently not authenticated) before being communicated to the untrusted storage device. As another example, the fastest signature systems involve

- hashing all the data being signed;
- some asymmetric work independent of the data length.

Similarly, a TLS session applies an authenticated cipher to every byte of data, after an asymmetric handshake independent of the data length. Trends toward larger data volumes, supported by faster CPUs and faster networks, mean that a larger and larger fraction of overall cryptographic cost comes from symmetric cryptography, providing one reason for having more symmetric competitions than asymmetric competitions.

Perhaps more advanced cryptographic operations will dominate cryptographic costs someday. The yearly iDASH “secure genome analysis competition” [128] measures performance of homomorphic encryption, multiparty computation, etc. As another example, to the extent that post-quantum cryptography is more expensive than pre-quantum cryptography, it changes the balance of costs—which, again, does not imply that its costs matter to the end users; this is something that needs analysis.

Comparing the symmetric competitions shows trends toward larger and more complex inputs and outputs in the cryptographic algorithm interfaces. DES has a 64-bit block size; AES has a 128-bit block size. Stream ciphers encrypt longer messages. Hash functions hash longer messages. Authenticated ciphers include authentication tags in ciphertexts, and optionally authenticate another input. Many NISTLWC submissions support hashing and authenticated encryption, sharing resources between these functions. The main driver of these trends is that symmetric algorithms with larger interfaces often reach levels of efficiency that seem hard to achieve with smaller interfaces, although there are also some security arguments for larger interfaces. See generally [22, Section 2].

2.6. *How AES Performance was Compared*

During the AES competition, Biham [37, Table 3] reported that “the speed of the candidate ciphers on Pentium 133MHz MMX” was 1254 cycles for Twofish, 1276 cycles for Rijndael, 1282 cycles for CRYPTON, 1436 cycles for RC6, 1600 cycles for MARS, 1800 cycles for Serpent, etc. for encrypting a 128-bit block under a 128-bit key.

Serpent, generally viewed as the AES runner-up, had (and has) a much larger security margin than Rijndael, the eventual AES winner. Biham also reported speeds scaled to “proposed minimal rounds”: 956 cycles for Serpent (17 rounds rather than 32), 1000 cycles for MARS (20 rounds rather than 32), 1021 cycles for Rijndael (8 rounds rather than 10), etc.

Let’s focus on Biham’s reported 1276 cycles for full 10-round Rijndael, almost 80 cycles per byte. The Pentium (with or without MMX) could run at most 2 instructions per cycle, and obviously it didn’t have AES instructions, but did it really need 1276 cycles for 160 table lookups and some auxiliary work? No, it didn’t. Schneier, Kelsey, Whiting, Wagner, Hall, and Ferguson [121, Table 2] reported Rijndael taking 320 Pentium cycles, just 20 cycles per byte. They also estimated that Serpent would take 1100 Pentium cycles—but then Osvik [105, page 8] reported an implementation taking just 800 cycles.

Compared to Biham’s reports, Serpent was more than $2\times$ faster, and Rijndael was $4\times$ faster. Overall these speedups seem to favor Rijndael, for example putting Rijndael

ahead of Serpent in the “proposed minimal rounds” speed. On the other hand, overall these speedups *compress* the difference in costs between Serpent and Rijndael, from $1800 - 1276 = 524$ cycles to $800 - 320 = 480$ cycles; and these speedups make it more likely that both ciphers will meet the users’ performance requirements.

Why did these reports end up with such different numbers? And why did NIST’s AES efficiency testing [99] feature CRYPTON as the fastest candidate in its tables and its graphs, 669 Pentium Pro cycles to encrypt, with Rijndael needing 809 Pentium Pro cycles to encrypt? The Pentium Pro is generally faster than the Pentium (and Pentium MMX); [121] reported 345 Pentium Pro cycles for CRYPTON and 291 Pentium Pro cycles for Rijndael.

2.7. The Process of Comparing Cryptographic Speeds

All of these speed numbers arise from the general process shown in Fig. 2. The first column has a cryptographic algorithm: for example, the Rijndael encryption algorithm mapping a 128-bit plaintext and a 128-bit key to a 128-bit ciphertext. The second column has a programmer writing software for this algorithm—or for another algorithm computing the same mathematical function. The third column has a benchmarking mechanism that measures the speed of the software, for example the number of cycles that the software uses on a 133MHz Pentium MMX CPU. The fourth column has an advertisement mechanism that might or might not bring the resulting cycle count to the attention of readers.

There are several reasons that the outputs of this process vary:

- The cryptographic functions vary: e.g., Rijndael and Serpent have different speeds. The whole point of a speed competition is to compare the speeds of different functions.
- The CPUs vary. This complicates comparisons. If function F is faster than function G on one CPU, but slower on another, then which function wins the competition? What happens if F is slower than G on today’s CPUs, but faster than G in hardware, perhaps hardware built into future CPUs?
- The software varies. A programmer often fails to achieve the best speed for a cryptographic function on a CPU. The slowdown depends on many details of the function, the CPU, the programmer’s experience, and the programmer’s level of enthusiasm for the function. There are many counterexamples to the notion that the slowdown is independent of the function: for example, compared to [121], NIST’s study [99] slowed down CRYPTON by a factor 1.94, slowed down Rijndael by a factor 2.78, and reversed the comparison between the algorithms.
- The benchmarking mechanism varies, for example in the handling of per-input timing variations, initial code-cache-miss slowdowns, operating-system interrupts, clock-frequency variations, and cycle-counting overheads.
- The advertisement mechanism varies. As an example, measurements that are slower than previous work are likely to be suppressed if the advertisement mechanism is a paper claiming to set new speed records, but less likely to be suppressed if the advertisement mechanism is a paper claiming to compare multiple options.

Both 1276 cycles from [37] and 320 cycles from [121] are reported to be Rijndael measurements on the Pentium, so the first and second effects cannot explain the gap. The easiest explanation is the third effect, although it is easy to imagine some contributions from the fourth and fifth effects.

The situation is different when cryptographic functions are deployed. The CPUs still vary, but, for each CPU, slower software is systematically suppressed in favor of faster software (as measured by a unified benchmarking mechanism), because users who care about speed don't want the slower software. For example, the OpenSSL cryptographic library contains 26 AES implementations, almost all in assembly language, in a few cases weaving AES computations together with common hash-function computations. The library checks the target platform and selects an implementation accordingly.

When different implementations run at different speeds for the same function on the *same* CPU, what speed does a speed competition assign to that function? Here are two strategies for answering this question:

- The **“fair and balanced” strategy** gives equal weight to the speeds of all implementations.
- The **real-world strategy** takes the fastest available implementation, the same way that a cryptographic library does, while suppressing the speeds of slower implementations.

As soon as there are two implementations running at different speeds, the “fair and balanced” strategy reports worse speed than the real-world strategy, speed farther from what the users will see (assuming the users care about speed). The real-world strategy creates a healthy incentive for implementors to look for and eliminate slowdowns in their own implementations, while the “fair and balanced” strategy creates an unhealthy incentive for implementors to “accidentally” create slow implementations of competing functions.

The standard argument for the “fair and balanced” strategy is to say that reducing CPU time is rarely worth the software development time. Knuth [86, page 268] famously expressed the trade-off as follows:

Programmers waste enormous amounts of time thinking about, or worrying about, the speed of noncritical parts of their programs, and these attempts at efficiency actually have a strong negative impact when debugging and maintenance are considered. We *should* forget about small efficiencies, say about 97% of the time; premature optimization is the root of all evil.

But Knuth's complaint here is about optimizing “noncritical parts” of programs. Knuth continued as follows:

Yet we should not pass up our opportunities in that critical 3%. A good programmer will not be lulled into complacency by such reasoning, he will be wise to look carefully at the critical code; but only after that code has been identified.

Today there are tens of millions of lines of code in a web browser, and many more lines of code in a complete computer system. Almost all of this code disappears if one asks which code has a noticeable impact on performance. A few “hot spots” are so important that implementors look carefully at making them run as quickly as possible.

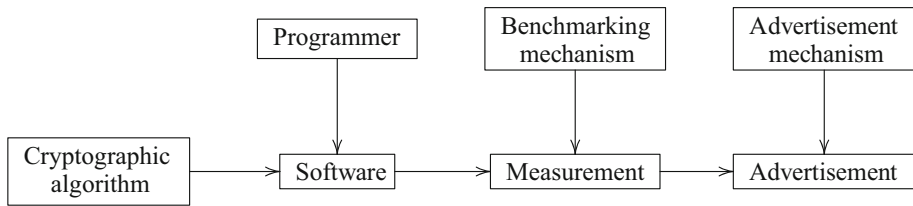


Fig. 2. The process of producing performance data for cryptographic software .

If a cryptographic operation is not one of these “hot spots,” then why is it the topic of a speed competition?

2.8. A Diagram

See Fig. 2.

2.9. How AES Speeds were Compared, Part 2

Say someone publishes a faster implementation of a cipher, 500 lines of software, 2 years after a cipher competition begins. Does this mean that the software took 24 person-months of work, and that similarly optimizing for the 20 most important platforms would take 480 person-months of work? Or could it be that the same person was busy with multiple projects, spending only 2 months of work on this project, and could it be that a new platform shares 75% of the optimization work with previous platforms, so optimizing for the 20 most important platforms would take under 12 person-months of work? Even if it’s really 480 person-months, wouldn’t the community invest this effort in any widely deployed cipher? Compare [91], which estimates that AES produced \$250 billion in worldwide economic benefits between 1996 and 2017.

NIST’s AES report [102] did not measure AES code-development time but claimed that this time was often important:

In some environments, the speed at which the code runs is perceived as a paramount consideration in evaluating efficiency, overriding cost considerations. In other cases, the time and/or cost of code development is a more important consideration.

NIST deviated slightly from the “fair and balanced” strategy, and, in particular, refused to list speeds of the fast Serpent implementations from [105] and [65], since those implementations had been constructed by “1000h of execution of search programs” and “do not necessarily port to different platforms.”

Around that time, the Internet was reported to be communicating roughly 2^{58} bytes per year, more than doubling every year. The load was spread across millions of CPUs. It is hard to see how a few dozen CPUs spending a day on “execution of search programs” can be a cost to worry about, even if one repeats those hours for dozens of different target platforms. Furthermore, it is easy to see that the optimizations from [105] and [65] work

reasonably well on a wide range of platforms, even if further searching would do better on some platforms.

It is important to realize the mismatch between the resources available for widely deployed algorithms and the resources available *during* a competition. The costs of cryptographic optimization—in human time or computer time—can be huge obstacles for submitters without serious optimization experience. The submitters ask for help from people with experience, but the only people paying attention at this point are crypto junkies, and there are many submissions. Any particular algorithm struggles to gain attention. It is easy to see how this struggle could be misinterpreted as a reflection of the deployment environment, rather than as a problem created by the competition process.

Think about the cryptographic function that would win a speed competition if it were properly optimized. There is a risk that this function loses the competition because the necessary optimizations are not demonstrated in time. Ways to reduce this risk include

- specifying a small number of target platforms as the arenas for competition, to better focus optimization work during the competition (although there is then a risk that these platforms are inadequate representatives of other platforms);
- tracking expert assessments of the unexplored avenues for speedups in each submission; and
- extending the competition time until the performance picture has settled down.

The “fair and balanced” strategy exacerbates this risk by assigning a low weight to speedups found later in the competition, whereas the real-world strategy ignores all worse speeds the moment that better speeds have been demonstrated.

One might think that Rijndael was much faster than Serpent even after the speedups, so assigning higher weights to the speedups could not have changed the AES selection. But the speed picture was not this simple. The hardware data surveyed in [102] suggested that the most efficient proposal for hardware encryption was pipelined counter-mode encryption using Serpent.⁵ Equalizing security margins would have made Serpent much faster. Rijndael was faster in software because its 256-entry table lookups reused existing CPU instructions—but hardware implementations have to pay for table lookups. The optimizations from [105] and [65] should have increased Serpent’s hardware advantage while decreasing its software disadvantage. If I’ve counted correctly then [65] uses 201 bit operations per plaintext bit for full 32-round Serpent (plus 1 bit operation per plaintext bit for counter-mode xor and a small cost for counter maintenance), very much like the best operation count known today for 10-round Rijndael.

2.10. Better Benchmarking Mechanisms

NESSIE, mentioned in Sect. 1, was a 2000–2003 EU project “New European Schemes for Signatures, Integrity, and Encryption.” I’m not sure NESSIE qualifies as a competition—it selected 17 algorithms—but in any case it took important steps toward matching its performance evaluations with the reality of what cryptographic users would see. NESSIE

⁵There were complaints about Serpent using extra hardware for the inverse function, but counter-mode encryption does not need the inverse.

published a software API supporting secret-key encryption, public-key signatures, etc.; collected C implementations of many cryptographic functions, with all implementations using the same API; tuned the C implementations for speed; wrote a benchmarking toolkit to measure speed; ran the benchmarking toolkit on many computers; and published the results. See [110].

As part of the eSTREAM competition, Christophe De Cannière developed a new API for stream-cipher software, and wrote a new benchmarking toolkit [40] to measure implementations supporting the API. This toolkit was limited to stream ciphers but had several advantages over the NESSIE toolkit. Notably, it tried more compiler options; it supported assembly-language software; and it was *published*. Implementors could run the toolkit to quickly and reliably see how fast their own software was, and to guide improvements to the software. Third parties could run the toolkit to contribute and verify public benchmark results. The quick feedback to implementors—from running the toolkit on their own machines and from seeing results announced by third parties—led to many implementation improvements during eSTREAM. The toolkit followed the real-world strategy, automatically reporting a list of ciphers with the speed of the fastest implementation of each cipher; see [40, Section 6].

In its final AES report [102, Section 2.5], NIST had complained that requests to consider a different number of rounds for an AES submission “would impact the large amount of performance analysis” that had already been done, since “performance data for the modified algorithm would need to be either estimated or performed again.” It wasn’t realistic to expect all the authors of performance-comparison papers to integrate new functions into their private benchmarking procedures and update their papers accordingly. This complaint goes away when the benchmarks come from an easily extensible *public* benchmarking toolkit: anyone can tweak the number of rounds in the implementations and run the toolkit again.

In 2006, Tanja Lange and I started eBATS, a new benchmarking project for public-key systems. Together with Christof Paar, Lange was the leader of the Virtual Application and Implementation Research Lab, VAMPIRE, within a European network, ECRYPT; the name “eBATS” stands for “ECRYPT Benchmarking of Asymmetric Systems.” STVL, ECRYPT’s Symmetric Techniques Virtual Lab, was running eSTREAM.

Lange and I designed a simple new cryptographic API to handle the needs of benchmarking *and* the needs of cryptographic libraries, so writing software for benchmarking was no longer inherently a separate task from writing software for real-world use. This increased the incentive for implementors to support the benchmarking API, and decreased the extra implementation effort. We analyzed and improved the end-to-end benchmarking process that turned new software into the public presentation of measurements. Extra feedback to implementors added extra incentives to contribute implementations.

In 2008, eSTREAM was drawing to a close, and the SHA-3 competition had been announced. Lange and I started eBACS, a unified benchmarking project that includes eBASC for continued benchmarking of stream ciphers, eBASH for benchmarking of hash functions, and eBATS. We replaced BATMAN, the original benchmarking toolkit for eBATS, with a new benchmarking toolkit, SUPERCOP. By late 2009, eBASH had

collected 180 implementations of 66 hash functions in 30 families.⁶ eBASH became the primary source of software performance information for the SHA-3 competition. See [45].

eBACS has continued since then, adding more cryptographic functions, more implementations of those functions, and newer CPUs—while continuing to run benchmarks on years of older CPUs for comparability. The SUPERCOP API was carefully extended to handle more operations, such as authenticated encryption. CAESAR, NISTPQC, and NISTLWC required submissions to provide software using the SUPERCOP API. SUPERCOP now includes 4350 implementations of 1369 cryptographic functions in hundreds of families. See [32].

This is not the end of the story. SUPERCOP measures cryptographic speeds on CPUs large enough to run Linux, but what about microcontrollers? FPGAs? ASICs? Performance metrics other than speed, such as energy usage? Notable efforts to improve benchmarking processes include the ongoing ATHENA project for FPGAs and ASICs, and the XBX, FELICS, XXBX, FELICS-AEAD, and `pcm4` projects for microcontrollers. See generally [64], [136], [137], [58], [42], [60], [83], and [82].

3. Security

Here we are, halfway through a paper that claims to be analyzing the extent to which competition procedures reduce security risks, and all I’ve been talking about is speed competitions. This section closes the gap.

3.1. *The Complex Relationship Between Speed and Security*

Let’s begin with the obvious argument that a cryptographic speed competition *is* a security competition.

Risk #1 of cryptography is that the cryptography isn’t used. One reason that cryptography isn’t used, as mentioned in Sect. 1 and illustrated by the 4-year delay in Android encryption reviewed in Sect. 2.1, is that the cryptography doesn’t meet the user’s speed requirements. Perhaps the requirements are driven by reality—something with worse efficiency would, if deployed, be a problem—or perhaps they are driven by *fear* that there will be a problem. Either way, something that doesn’t meet the requirements won’t be deployed, and if nothing meets these requirements then nothing will be deployed. A cryptographic speed competition identifies the functions that have the best chance of meeting the user’s speed requirements.

There is, however, an equally obvious argument in the opposite direction, namely that a cryptographic speed competition naturally identifies the *weakest* functions. RSA-1024 is more efficient than RSA-2048, and RSA-512 is more efficient than RSA-1024, so RSA-512 will beat RSA-1024 and RSA-2048 in a speed competition—but RSA-512 is breakable. AES candidates were slower with 256-bit keys than with 128-bit keys.

⁶I don’t mean to suggest that “family” has a clear definition here. Is SHA-1 in the same family as SHA-224, SHA-256, SHA-384, and SHA-512? The benchmarking process measures each function separately.

Rijndael proposed not just multiple key sizes but also multiple block sizes, with top speed requiring the minimum block size. DES had just one version, but Diffie and Hellman proposed a longer-key variant, and people complained that this was more expensive; see Sect. 2.5.

If each family of algorithms claims a trade-off between security and efficiency, then it is unsurprising that graphing the claimed security and efficiency of many different proposals will also show such a trade-off; see, e.g., the general slant of the graphs in [24]. How, then, is a competition for top speed not the same as a competition for minimum security? The users will have something that meets their performance requirements—something breakable.

Most competitions respond by specifying a **minimum allowed security level**. A more subtle extension of this response is to say that users should take the **maximum security margin**. This works as follows:

- Identify the most efficient cryptographic function that seems to meet or exceed the minimum allowed security level. This is a speed competition, but subject to a security requirement.
- Within the family containing the most efficient function, take the *largest* function that meets the users' performance requirements.

The idea here is that the speed competition gives users the maximum room for larger keys, more cipher rounds, and other forms of security margins that—we hope—provide a buffer against attack improvements.

For example, say the efficiency metric is bit operations per bit of plaintext to encrypt a long stream; and say the minimum allowed security level is 2^{128} . My understanding of current attacks is that Serpent reaches this security level with 12 rounds, using about 75 operations per bit; Rijndael reaches this security level with 8 rounds, using about 160 operations per bit; and Salsa20 reaches this security level with 8 rounds, using 54 operations per bit. If these are the competitors then Salsa20 wins the speed competition.⁷ A user who can afford, say, 80 operations per bit then takes 12 rounds of Salsa20 (78 operations per bit). The same user would also be able to afford 12 rounds of Serpent, but 12 rounds of Salsa20 provide a larger security margin, presumably translating into a lower risk of attack.

The reality, however, is that cryptographic designers are overconfident, and see negligible value in large security margins (“I can’t have missed something so big”), even when history shows one example after another of attacks that would have been stopped by large security margins. Meanwhile the same designers see that proposing something with a large security margin is risky. Serpent proposed more than twice as many rounds as necessary and had the whole proposal dismissed as being too slow.

I knew that Salsa20 had far more rounds than I could break, correctly guessed that it had far more rounds than anyone else could break, correctly guessed that it would be competitive in speed anyway, and concluded that it would be able to get away with a large security margin. At the same time, knowing what had happened with Serpent, I

⁷This is an unfair comparison. Salsa20 was designed years later, taking advantage of lessons learned from Serpent and many other designs. Salsa20 also benefits from spreading differences through a 512-bit block, while the AES competition required 128-bit blocks and discarded Rijndael’s 256-bit-block options.

didn't want to go beyond 20 rounds. I held off on proposing reduced-round versions: an initial proposal with (say) 12-round and 20-round options would have been interpreted as indicating a lack of confidence in 12 rounds, whereas an initial proposal of 20 rounds followed by "Look, people can't break 12 rounds, and can't even break 8 rounds" sounded purely positive. There was an eSTREAM requirement to support 128-bit keys, but I proposed as many rounds for 128-bit keys as for 256-bit keys, so that users wouldn't have a speed incentive to take smaller keys.

All of this was converting a presumed software-speed advantage into extra security margin—but if someone else had designed a similar stream cipher with a smaller round-count parameter then Salsa20 would have been eliminated. In the opposite direction, there is a long history of submissions being eliminated from competitions because they chose parameters *slightly* too small for security—even if larger parameters would have been competitive. Is a competition supposed to be evaluating the best trade-offs available between speed and security, or is it supposed to be evaluating the designer's luck in initial parameter selection?

Submitters see what happened in previous competitions. This feedback loop keeps most security margins in a narrow range. Designers and implementors don't want to risk making mistakes in providing larger options. In the end, users aren't being given the choice to take the largest security margin they can afford. I did convince the relevant people that TLS would be just fine in performance using 20 rounds of ChaCha rather than 12, but most deployed security margins are much smaller than this, and competitions follow suit.

3.2. *The Complex Relationship Between Speed and Security, Part 2: Later Discovery of Attacks*

If we've correctly evaluated the security level of a cryptographic algorithm, and if that security level is high enough compared to the resources available to the attacker, then we shouldn't need a security margin—the algorithm is secure. The basic problem here is that we don't have procedures to reliably evaluate the security level of a cryptographic algorithm. Sometimes breaking an algorithm takes years or even decades of public attack development. This does not mean that the algorithm was secure in the meantime: the algorithm was never secure, and large-scale attackers could have found the break long before the public did.

MD5 was published in 1992 [112] with a conjecture of 2^{64} collision security. There were alarm bells from cryptanalysts, such as [59] ("it is anticipated that these techniques can be used to produce collisions for MD5"), but the conjecture was not publicly broken until 12 years later, when the results of [134] were announced. Follow-up work culminating in [126] exploited MD5 chosen-prefix collisions to efficiently forge certificates for arbitrary web sites. It was announced in 2012 that malware called "Flame" had been exploiting MD5 collisions since at least 2010; the analysis of [125] concluded that the Flame attackers had used an "entirely new and unknown" variant of [126] (meaning new from the public perspective), that the Flame design "required world-class cryptanalysis," and that it was "not unreasonable to assume" that this cryptanalysis predated [126].

As another example, SIDH appeared in 2011 [80] with a claim to improve upon "all previous isogeny-based schemes by orders of magnitude in performance at conventional

security levels.” SIDH became the foundation of the SIKE submission to NISTPQC, and its performance attracted interest. SIKE was used in a large-scale experiment [88] by Cloudflare and Google, was advertised in [49] as “a decade unscathed,” and was one of just four algorithms that [10] selected for consideration in the fourth round of NISTPQC (“SIKE remains an attractive candidate for standardization because of its small key and ciphertext sizes”), on top of the four algorithms selected for standardization at the end of the third round. SIKE was then broken by [43] and, independently, [94]. (For earlier security warnings, see [31, minute 48:25] and [27].) The largest proposed SIKE parameters, designed for 2^{256} security, were broken in [56] in 11 seconds. The same attacks do not seem to apply to the older isogeny-based system from [118].

Think of a cryptographic proposal as a random variable, with some probability $p(M)$ of being publicly broken within M months (i.e., there is probability $p(M)$ that a break is published $\leq M$ months after the proposal). Assume for simplicity that these probabilities are independent across proposals. Let’s also optimistically assume that $p(\infty)$, the limit of $p(M)$ as M increases, captures all attacks that the attacker will find—the public will eventually find everything; and that the same probabilities apply specifically to submissions to competitions, rather than competitions tending to encourage weak submissions.

By collecting enough data, we can retrospectively estimate $p(M)$ for small values of M , and perhaps the curve would let us guess $p(\infty)$. For example, one can start by estimating $p(12) \approx 1/3$ given that 5 of the 15 AES submissions were publicly broken within 12 months, although obviously this selection of data should be replaced by much more data.

Consider a competition that chooses a random winner among the submissions not publicly broken within 36 months (assuming there is one). A submission has

- probability $p(36)$ of being publicly broken within 36 months,
- probability $p(\infty) - p(36)$ of being breakable but not publicly broken within 36 months, and
- probability $1 - p(\infty)$ of being secure.

The winner is thus breakable with probability $(p(\infty) - p(36))/(1 - p(36))$.

If, for example, data collection shows that there have been 1000 cryptographic proposals, with just 100 publicly broken within 36 months, and just 10 (like MD5) publicly broken between 36 months and 180 months, then $1 - p(36) = 0.9$ and

$$p(\infty) - p(36) \geq p(180) - p(36) = 0.01,$$

so the winner is breakable with probability at least $0.01/0.9$. Does it make sense for cryptographers to worry about a user choosing a weak key with probability 2^{-64} , while not obviously worrying about each new cryptographic competition choosing a breakable winner with probability above 1%?

Now let’s partition the proposals into two types, 50% faster proposals and 50% slower proposals. Let’s define $p_1(M)$ as the conditional probability of a faster proposal being publicly broken within M months, and $p_2(M)$ as the conditional probability of a slower proposal being publicly broken within M months. By definition $p(M) = 0.5p_1(M) +$

$0.5p_2(M)$. Assume for simplicity that there are no further correlations between efficiency and brokenness.

Consider a competition that receives F faster submissions, receives infinitely many slower submissions, and chooses the *most efficient* submission not publicly broken within 36 months. The probability that all of the faster submissions are publicly broken within 36 months is $p_1(36)^F$, so the competition winner is breakable with probability

$$p_1(36)^F \frac{p_2(\infty) - p_2(36)}{1 - p_2(36)} + (1 - p_1(36)^F) \frac{p_1(\infty) - p_1(36)}{1 - p_1(36)}.$$

If $p_1(36)$ isn't too close to 1 and F isn't too small then $p_1(36)^F$ is close to 0, so the winner is breakable with probability close to $(p_1(\infty) - p_1(36))/(1 - p_1(36))$.

This probability could be even larger than $(p(\infty) - p(36))/(1 - p(36))$, meaning that taking the most efficient unbroken submission is increasing risk compared to taking a random unbroken submission. It's easy to imagine reasons for faster proposals to be more likely to be broken than slower proposals—and more likely to be publicly broken after 36 months, as in the case of MD5. On the other hand, perhaps data collection will show that for some reason faster proposals are actually less risky overall. Even if they're more risky, perhaps this is outweighed by the benefit that they produce *for users taking the maximum security margin*. Similar comments apply when there are more than 2 different levels of efficiency.

Perhaps $p(M)$ and $p_1(M)$ should be stratified into $p(M, Y)$ and $p_1(M, Y)$, where Y is the year when a proposal was made. Optimists might hope that $p(M, Y)$ has been decreasing with Y . But many submissions to the most recent competitions have been broken (see, e.g., [25, PDF page 91]), showing that $p(M, Y)$ remains far above 0 for small M . Why shouldn't we think that $p(\infty, Y)$ is even larger than $p(36, Y)$, and that $p_1(\infty, Y)$ is even larger than $p_1(36, Y)$?

3.3. The Overworked Cryptanalyst

One way to argue that competitions reduce security risks is to argue that they focus the community's attention on finding all possible attacks. But does a competition provide enough time for this?

F-FCSR, one of the eight ciphers selected for the eSTREAM portfolio, was then shown in [72] to be very efficiently broken from 13 megabytes of output. As another example, the Rijndael designers and NIST claimed that Rijndael was “not vulnerable to timing attacks,” but this was then disproven; as noted in Sect. 1, timing attacks continue to cause AES security problems today. These attacks work for any number of AES rounds, illustrating that security margins don't necessarily eliminate the security risks that remain after competitions.

If an attack takes years to develop, perhaps the reason is that it is the end of a long sequential chain of thoughts adding up to years of latency, but a simpler explanation is throughput. The world has a limited number of cryptographic experts capable of carrying out, and willing to carry out, public security analysis. This valuable time is divided across a huge number of proposals. This problem is more severe for competitions having more submissions—and for competitions having more complicated submissions. A competition might attract cryptanalyst time that would otherwise have been spent

elsewhere, but the existence of a competition also *creates* submissions, cryptographic proposals that would not have existed otherwise, so it is far from clear that the amount of analysis per year per submission is larger than the amount of analysis per year per proposal outside competitions.

Perhaps having a critical mass of cryptanalysts focusing on one topic at the same time leads to breakthroughs that would not otherwise happen. Or perhaps the focus is wasting valuable cryptanalytic time on redundant parallel work, and cryptanalysts work more efficiently when they are roaming free through a larger field of targets.

Competitions normally run through multiple phases, each phase narrowing the list of submissions. See Fig. 3. A shorter and shorter list makes it more and more reasonable to believe that cryptanalysts are focusing on each remaining proposal more than what would have happened without a competition. AES, for example, narrowed the 15 initial submissions to 5 finalists, and presumably those were the top targets for security analysis at that point. On the other hand, final comments were due just 9 months later, and NIST announced the winner just 5 months after that. Concerns about not having enough time were expressed in [109, Section 2.1]:

We believe that the effort spent on evaluating the security of the five AES finalists has been very limited, certainly compared to the 17 man-years spent by IBM on DES in the 1970s.

Sometimes competitions ask cryptanalysts to focus on particular submissions, while not necessarily excluding other submissions:

- eSTREAM’s second phase selected 27 submissions, designating 10 as “focus” submissions: “These are designs that eSTREAM finds of particular interest. We particularly encourage more cryptanalysis and performance evaluation on these primitives.” The others were “designs that eSTREAM wishes to move to the second phase of the eSTREAM project.” (Two of the others, F-FCSR and Rabbit, ended up being in the eSTREAM portfolio.)
- NISTPQC’s third phase selected 15 submissions, designating 7 as finalists: “NIST intends to select a small number of the finalists for standardization at the end of the third round. In addition, NIST expects to standardize a small number of the alternate candidates (most likely at a later date).”

Cryptanalysts can also decide on their own to focus on the fastest submissions, guessing that those are the easiest submissions to break, and the most likely to be selected if they are not broken. This might seem to be a successful strategy if the focus produces an attack—but why should we think that enough time has been spent analyzing the remaining submissions?

Section 2.9 considered ways to reduce the risk of the fastest function not being recognized as the fastest. One can analogously try to reduce the risk of the fastest function not being recognized as being breakable:

- Limit the complexity of the security goals for the competition, to better focus security analysis work during the competition (although there is then a risk that these security goals are inadequate representatives of other security goals).
- Track expert assessments of the unexplored avenues of attack against each submission.

- Extend the competition time until the attack picture has settled down.

I suspect that collecting historical data will show that the security risks from later attack improvements have been quantitatively more severe, in probability and in impact, than the security risks arising from later performance improvements.

At the beginning of 2012, the SHA-3 competition was almost over, and the consensus of the cryptanalysts and designers I talked to was that it would be useful to have a new competition. Authenticated encryption was an obvious target—compared to stream ciphers and hash functions, authenticated ciphers are a step closer to the typical user’s needs—and group discussions didn’t identify any better targets. Interfaces even closer to the user’s needs were identified (for example, secure sessions) but raised concerns of being too big a jump, needing too much new security analysis.

When I announced CAESAR at the beginning of 2013, I posted a “Timeline (tentative)” stretching 5 years into the future, with a submission deadline a year later and then four years of analysis ending with a portfolio. The actual schedule ended up lasting a year longer than the tentative schedule: submitters were asking for more time already from the outset (e.g., “the extra 2 months is a small price to pay if it increases the quality of submission pool (which I’m sure it will)”), cryptanalysts were asking for more time, etc. In the end CAESAR selected a portfolio of authenticated ciphers with exciting performance features and security features—see Sect. 2.5 for the list of ciphers—and *in the subsequent 43 months* nothing has publicly gone wrong with any of them.

The italicized phrase is important. There is still no guarantee that enough time has been taken to find the best attacks against these ciphers. It is concerning to see, e.g., the breaks in [77] and [93] of various versions of OCB (OCB2 for [77], and OCB3 with nonces shorter than 6 bits for [93]), even though those are not exactly the versions of OCB submitted to CAESAR.

3.4. Timelines

See Fig. 3.

3.5. Cryptographic Risk Management Beyond Timing

Public advances in attack algorithms generally begin with experts recognizing dangerous structure in the functions being attacked. Typically this structure can be described as an analogy to another broken function.

This doesn’t mean that an expert recognizing dangerous structure is always able to find an attack. Often one cryptanalyst extends an attack on function F to an attack on function G , and then another cryptanalyst extends the attack on G to an attack on function H , where the first cryptanalyst already recognized the analogy between F and H and figured out the attack on G as one step from F toward H . Sometimes the first cryptanalyst already issues a warning regarding H , such as the MD5 alarm bells from [59] mentioned in Sect. 3.2.

A competition doesn’t have to select the most efficient unbroken submission. It can try to reduce security risks by paying attention to extra information, namely the concerns

competition	years	
DES: the Data Encryption Standard	1974–1976	
AES: the Advanced Encryption Standard	1998–2000	
eSTREAM: the ECRYPT Stream Cipher Project	2005–2008	
SHA-3: a Secure Hash Algorithm	2008–2012	
CAESAR: Competition for Authenticated Encryption: Security, Applicability, and Robustness	2014–2019	
NISTPQC: NIST Post-Quantum Cryptography Standardization Project	2017–?	
NISTLWC: NIST Lightweight Cryptography Standardization Project	2019–2023	

competition	start	submissions
DES	M–13	M0: 1 → M26: 1
AES	M–17	M0: 15 → M14: 5 → M28: 1
eSTREAM	M–5	M0: 34 → M11: 27 → M24: 16 → M36: 8
SHA-3	M–21	M0: 51 → M9: 14 → M26: 5 → M48: 1
CAESAR	M–14	M0: 56 → M16: 29 → M29: 15 → M48: 7 → M59: 6
NISTPQC	M–20	M0: 69 → M13: 26 → M31: 15 → M55: 4+4 → ?
NISTLWC	M–30	M0: 56 → M6: 32 → M25: 10 → M48: 1

Fig. 3. Number of submissions remaining in each phase of various competitions. M0 is the calendar month when initial submissions were due. The first boldface number is the number of submissions allowed into the first phase. This is often smaller than the total number of submissions; e.g., DES disallowed most submissions, according to [55, Section 6]. Each subsequent M is the calendar month when submissions were announced for a subsequent phase, and the boldface number is the number of those submissions. For DES, AES, eSTREAM, SHA-3, CAESAR, and NISTLWC, the final boldface number is the number selected for the portfolio as output of the competition, ending the competition. eSTREAM then updated its portfolio 5 months later to remove a broken portfolio member. NISTPQC is ongoing, with **4+4** meaning that 4 candidates were announced as being selected for the portfolio and 4 candidates were announced as being under continued consideration for the portfolio. Second column is when competition was announced. Range of years in the top table starts when initial submissions were due and ends when the portfolio was announced .

that experts have regarding submissions. Factoring this information into decisions is complementary to giving cryptanalysts more time, the approach of Sect. 3.3.

I’m not saying that analogies always turn into attacks. It’s easy to draw a chain of analogies from any cryptographic function to a broken function, so if analogies always turned into attacks then everything would be broken, which we hope isn’t the case. There is also a procedural problem with simply downgrading any submission for which someone claims that there’s a dangerous structure: this invites superficial claims from competing submissions.

I established a general policy of public evaluations for CAESAR:

CAESAR selection decisions will be made on the basis of *published* analyses. If submitters disagree with published analyses then they are expected to promptly and *publicly* respond to those analyses. Any attempt to privately lobby the selection-committee members is contrary to the principles of public evaluation and should be expected to lead to disqualification.

Perhaps we can find clear rules reducing cryptographic risks, improving upon the baseline rule of eliminating publicly broken algorithms. Those rules can then be published, shown through analyses to be beneficial, and applied by everybody. But what happens when

experts are spotting risks in a way that isn't captured by any known rules? Do we ignore those risks, or do we try to take them into account?

One answer is to put the experts onto the selection committee. I tried hard to fill the CAESAR selection committee with top symmetric cryptographers, people having the experience and judgment to see risks in advance. I promised, as part of inviting people to join the committee and as part of the public procedures for CAESAR, that the committee would simply select algorithms and cite public analyses, rather than publishing its own analyses. Forcing the committee to publish analyses would have discouraged participation,⁸ taking resources away from the core job of making judgment calls *beyond* published analyses.

Almost everyone I invited said yes. A few later ran out of time, but all of the following continued through the end (affiliations listed here are from when CAESAR began):

- Steve Babbage (Vodafone Group, UK)
- Alex Biryukov (University of Luxembourg, Luxembourg)
- Anne Canteaut (Inria Paris-Rocquencourt, France)
- Carlos Cid (Royal Holloway, University of London, UK)
- Joan Daemen (STMicroelectronics, Belgium)
- Orr Dunkelman (University of Haifa, Israel)
- Henri Gilbert (ANSSI, France)
- Tetsu Iwata (Nagoya University, Japan)
- Stefan Lucks (Bauhaus-Universität Weimar, Germany)
- Willi Meier (FHNW, Switzerland)
- Bart Preneel (COSIC, KU Leuven, Belgium)
- Vincent Rijmen (KU Leuven, Belgium)
- Matt Robshaw (Impinj, USA)
- Phillip Rogaway (University of California at Davis, USA)
- Greg Rose (Qualcomm, USA)
- Serge Vaudenay (EPFL, Switzerland)
- Hongjun Wu (Nanyang Technological University, Singapore)

I served on the committee as non-voting secretary, tracking the discussions and decisions and handling public communication.

I don't know how to prove that factoring in expert judgments is more reliable than simply taking the fastest unbroken algorithm. Maybe it isn't—or maybe there's a better approach. It would be beneficial for the cryptographic community to put more effort into analyzing and optimizing risk management techniques.

⁸Consider an expert whose mental model says that unbroken algorithms A , B , C , D , and E have, respectively, chance 90%, 20%, 10%, 80%, and 60% of being secure, and consider a process that factors this information into decisions. The same expert sees ample evidence of the general public's limited understanding of what probabilities mean, and does not want to be subjected to complaints such as " E was broken after you said it was probably secure!".

3.6. The Goal of Limiting Security

Performance pressures and limited time for security analysis are not the only sources of security risks in cryptographic competitions, as the history of DES illustrates.

According to a 1978 interview [85], DES product leader Walter Tuchman described DES as “the culmination of six years of research and development at IBM,” a “three-pronged effort” involving his data-security-products group at IBM, the “mathematics department at IBM’s Yorktown Heights research center,” and “university consultants.” IBM was then “ready to respond” when the National Bureau of Standards (NBS, later renamed NIST) “issued its request for data encryption algorithm proposals.” Regarding accusations that IBM and NSA had “conspired,” Tuchman said “We developed the DES algorithm entirely within IBM using IBMers. The NSA did not dictate a single wire!”

In 1979, NSA director Bobby Inman gave a public speech [76] including the following comments: “First, let me set the record straight on some recent history. NSA has been accused of intervening in the development of the DES and of tampering with the standard so as to weaken it cryptographically. This allegation is totally false.” Inman continued with the following quote from a public 1978 Senate report [123]: “NSA did not tamper with the design of the algorithm in any way. IBM invented and designed the algorithm, made all pertinent decisions regarding it, and concurred that the agreed upon key size was more than adequate for all commercial applications for which the DES was intended.” This report was also mentioned in [85], which said that according to Tuchman the report had concluded “that there had been no collusion between IBM and the NSA.”

However, an internal NSA history book “American cryptology during the cold war” tells a story [81, pages 232–233] of much heavier NSA involvement in DES:

- NBS began to investigate encryption in 1968. NBS “went to NSA for help.”
- NSA’s “decision to get involved” with NBS on this was “hardly unanimous.” A “competent industry standard” could “spread into undesirable areas” such as “Third World government communications” and drugs and terrorism. On the other hand, “NSA had only recently discovered the large-scale Soviet pilfering of information from U.S. government and defense industry telephone communications. This argued the opposite case—that, as Frank Rowlett had contended since World War II, in the long run it was more important to secure one’s own communications than to exploit those of the enemy.”
- “Once that decision had been made, the debate turned to the issue of minimizing the damage. **Narrowing the encryption problem to a single, influential algorithm might drive out competitors, and that would reduce the field that NSA had to be concerned about. Could a public encryption standard be made secure enough to protect against everything but a massive brute force attack, but weak enough to still permit an attack of some nature using very sophisticated (and expensive) techniques?**” (Emphasis added. It is interesting to note the lack of any consideration of the possibility that any cryptosystem weak enough to be breakable by NSA would also be breakable by the Soviets.)
- Back to NBS: It “was decided” that NBS would “use the *Federal Register* to solicit the commercial sector for an encryption algorithm.” NSA would “evaluate the quality, and if nothing acceptable appeared, would devise one itself.”

- The response to NBS’s 1973 call for proposals “was disappointing, so NSA began working on its own algorithm.” NSA then “discovered that Walter Tuchman of IBM was working on a modification to Lucifer for general use. **NSA gave Tuchman a clearance and brought him in to work jointly with the Agency on his Lucifer modification.**” (Emphasis added.)
- Regarding the goal of making sure DES was “strong enough” but also “weak enough”: “NSA worked closely with IBM to strengthen the algorithm against all except brute force attacks and to strengthen substitution tables, called S-boxes. Conversely, NSA tried to convince IBM to reduce the length of the key from 64 to 48 bits. Ultimately, they compromised on a 56-bit key.” (For comparison, the Senate report had stated that “NSA convinced IBM that a reduced key size was sufficient” and that NSA had “indirectly assisted in the development of the S box structures.”)
- “The relationship between NSA and NBS was very close. NSA scientists working the problem crossed back and forth between the two agencies, and **NSA unquestionably exercised an influential role in the algorithm.**” (Emphasis added.)

The relevant portions of this book became public because, starting in 2006, the non-profit National Security Archive (abbreviated “The Archive,” not “NSA”) filed a series of declassification requests and appeals [101] regarding the book. This forced portions of the book to be released in 2008, and further portions to be released in 2013—officially documenting, for example, NSA surveillance of Martin Luther King, Jr., journalist Tom Wicker, and senator Frank Church. There were also some intermediate releases from the book in response to a FOIA request filed by John Young in 2009; see [142].

To summarize, NSA worked with NBS on the DES competition before the competition was announced, and worked jointly with IBM on the design of DES before the final design was submitted to the competition. NSA’s actual goals for the competition—goals that it acted upon, with considerable success—included (1) making sure that DES was “weak enough” to be breakable by NSA and (2) having DES be “influential” enough to “drive out competitors.” The first goal directly threatens security, and the second goal extends the security damage.

3.7. *The Difficulty of Recognizing Attackers*

An obvious response to the type of attack described in Sect. 3.6 is to set up competition procedures that exclude NSA—and other known attackers—from participation. However, NSA can secretly hire consultants to participate in the competitions and to try to weaken security in the same way that NSA would have. These consultants can deny NSA involvement, the same way Tuchman did.

The core problem is that it is not easy to recognize attackers. It is instructive to look back at the extent to which the cryptographic community has failed to recognize NSA as an attacker, never mind the harder problem of recognizing others working with NSA as attackers.

After differential cryptanalysis was published but was shown to have less impact on DES than on many DES variants, Coppersmith revealed [48] that the DES design team had already known about differential cryptanalysis and had designed DES accordingly.

This differential-cryptanalysis story contributed to a pervasive “good guys” narrative claiming that NSA had *strengthened* IBM’s DES design. Here are two examples of this narrative appearing in response to concerns regarding NSA influence:

- NIST’s standard elliptic curves were designed by NSA, and were claimed to be “verifiably random.” Scott [122] pointed out that if NSA knew a weakness in one curve in a million then the claimed “verifiable randomness” would not have stopped NSA from selecting a weak curve; see also [29], [28], and [30]. For a “good guys” response, see [68], which, in reply to [29], stated the following: “Flipside: What if NIST/NSA know a weakness in 1/10000000 curves? NIST searches space for curves that *aren’t* vulnerable.” (The same author later stated [69] that this comment was from when he was “younger and more naive.”)
- NSA budget documents leaked in September 2013 listed 0.25 billion dollars per year for a “SIGINT Enabling Project” that “actively engages the U.S. and foreign IT industries to covertly influence and/or overtly leverage their commercial products’ designs” to make them “exploitable” [107], including a goal to “influence policies, standards and specifications for commercial public key technologies.” This is what one would expect from an agency whose primary mission has always been signals intelligence; and it is consistent with NSA’s early-1970s goal, quoted above, of ensuring that DES was “weak enough to still permit an attack of some nature.” For a “good guys” response, see [39], which portrays the “SIGINT Enabling Project” as something new: [39] has subtitle “Leaked documents say that the NSA has compromised encryption specs. It wasn’t always this way”; claims that NSA’s “secretive work on DES” had “made the algorithm better”; and asks if there was a “change in mission.”

See also [78, 143], and [138].

The disclosures in 2013 did not stop NSA from participating in processes to select cryptographic algorithms. See, e.g., [14], describing NSA’s efforts between 2014 and 2018 to convince ISO to standardize Simon and Speck. One can only guess how many more algorithm-selection processes NSA was influencing through proxies in the meantime.

CAESAR began before those disclosures, but I was already well aware of NSA’s role as an attacker; see, e.g., [18]. I hoped that having as much as possible done in public would, beyond its basic advantages in correcting errors, also stop NSA and other attackers from sabotaging the process. Obviously attackers could still submit algorithms, but one of the required sections of each submission was the following:

Design rationale: An explanation of the choices made in the cipher design. This section is expected to include an analysis of how a weakness could be hidden in the cipher. This section must include the following statement: “The designer/designers have not hidden any weaknesses in this cipher.”

An attacker can easily lie about the existence of weaknesses, but being forced to explain the choices made in the design gives the community and the committee a chance to catch inadequate explanations.

Would an attacker be able to sneak a weak algorithm through a committee full of experts? Would it be able to sneak a weak algorithm through a competition that simply takes the fastest unbroken algorithm? These are interesting questions to analyze.

3.8. *The Goal of Producing Publications*

I'll close this paper by describing one more incentive that creates security risks in cryptographic competitions, an incentive that also explains many phenomena described earlier in this paper.

As academic cryptographers, we're paid primarily to produce publications, specifically papers. Most—although not all—deployed systems come from papers written by academic cryptographers, after passing through a long supply chain involving people paid to produce cryptographic standards, and people paid to produce cryptographic libraries, and so on.

When a cryptographic system fails, we blame *that system* for failing, as noted in Sect. 1. We then use the failure as motivation to write papers proposing and analyzing new systems. If the broken system is important enough then this also means new versions of standards, new versions of libraries, etc.

We all have a perverse incentive to stay in this situation, collectively creating a never-ending series of cryptosystems failing in supposedly new and exciting ways, so that we can continue writing papers designing and analyzing the next systems. Papers and grant proposals on improved attacks and improved cryptosystems and security proofs habitually explain their importance by citing recent failures. If we instead give the users “boring crypto” [20]—“crypto that simply works, solidly resists attacks, never needs any upgrades”—then will our readers and funding agencies still be interested in our subsequent papers? Perhaps, but do we really want to take this chance?

If we have boring block ciphers then as a community we could move on to, say, stream ciphers, and if we have boring stream ciphers then we could move on to authenticated ciphers, and if we have boring authenticated ciphers then we could move on to secure sessions. But won't the users say at some point that they have exactly the right cryptographic operations and don't need further input from us? It's safer for us if our core cryptography keeps failing.

The cryptographic community as a whole systematically flunks Taleb's “skin in the game” requirement [127] for risk management. As cryptographers, how often do we think that the damage caused by cryptographic failures will make *us* suffer? The designers of a system don't expect it to be broken in the first place. If it *is* broken then hey, look, the designers have another citation, and now we can all write follow-up papers. It's against community standards to blame the designers rather than blaming the broken system. At worst there's a brief moment of embarrassment if the attack was “too easy.”

We put *some* sort of requirements on cryptosystems to control the size of the literature and maintain the prestige of publications—e.g., any new block cipher must identify some performance metric where the cipher outperforms previous ciphers, and must include certain types of cryptanalysis—but we have little incentive to match these requirements to what the users want. What the users want, most importantly, is for us to be super-careful about security, but we have personal and community incentives against this. Being more careful than whatever is required for a publication is taking time away from

writing more papers, and as a community we want a sufficiently steady stream of broken cryptosystems as continued fuel for the fire.

Imagine a competition requiring every cipher to have twice as many rounds as it seems to need. This would make typical attack improvements less scary, and would eliminate most—although not all—cipher breaks. This would make papers harder to publish. The community thus has an incentive to argue *against* such a requirement, claiming that it's overkill and claiming that a $2\times$ slowdown is a problem. To support such arguments, we generate even more papers, such as performance-analysis papers saying that Serpent is slower than Rijndael.

More broadly, performance seems to be the most powerful weapon we have in the fight against ideas for reducing security risks. Performance constantly drives us toward the edge of disaster, and that's what we want. The edge is interesting. The edge produces papers. This also produces an incentive for us to continually claim that performance matters, and an incentive for us to avoid investigating the extent to which this is true. See Sect. 2.

A traditional report on the CAESAR competition would say that it produced many papers, advancing researchers' understanding of security and performance, building a foundation for the next generation of papers on symmetric cryptology. All of these things are true. DES was already a success in these metrics, and subsequent competitions have been even more successful. The challenge for the community is to figure out whether we can maintain success in what we're paid to do *without* a never-ending series of security failures.

Funding Open Access funding enabled and organized by Projekt DEAL.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- [1] – (no editor), Redacted transcript of hearing of House International Relations Committee on 21 July 1997 (1997). <https://cryptome.org/jya/hirhear.htm>.
- [2] – (no editor), International conference on field programmable logic and applications, FPL 2010, August 31 2010–September 2, 2010, Milano, Italy, IEEE Computer Society, 2010. ISBN 978-0-7695-4179-2.
- [3] – (no editor), 27th annual network and distributed system security symposium, NDSS 2020, San Diego, California, USA, February 23–26, 2020, The Internet Society, 2020. ISBN 1-891562-61-4.
- [4] – (no editor), 2020 IEEE Symposium on Security and Privacy (S&P 2020), 17–21 May 2020, San Francisco, California, USA, Institute of Electrical and Electronics Engineers, 2020. ISBN 978-1-7281-3497-0.
- [5] – (no editor), Proceedings of the 29th USENIX security symposium, August 12–14, 2020, USENIX, 2020.

- [6] Akamai, Akamai online retail performance report: Milliseconds are critical (2017). <https://web.archive.org/web/20170810113428/https://www.akamai.com/uk/en/about/news/press/2017-press/akamai-releasesspring-2017-state-of-online-retail-performance-report.jsp>.
- [7] Akamai, The state of online retail performance (2017). <https://web.archive.org/web/20200416173528/https://akamai.pdf>.
- [8] Gorjan Alagic, Jacob Alperin-Sheriff, Daniel Apon, David Cooper, Quynh Dang, Yi-Kai Liu, Carl Miller, Dustin Moody, Rene Peralta, Ray Perlner, Angela Robinson, Daniel Smith-Tone, Status report on the first round of the NIST Post-Quantum Cryptography Standardization Process, NISTIR 8240 (2019). <https://csrc.nist.gov/publications/detail/nistir/8240/final>.
- [9] Gorjan Alagic, Jacob Alperin-Sheriff, Daniel Apon, David Cooper, Quynh Dang, Yi-Kai Liu, Carl Miller, Dustin Moody, Rene Peralta, Ray Perlner, Angela Robinson, Daniel Smith-Tone, Status report on the second round of the NIST Post-Quantum Cryptography Standardization Process, NISTIR 8309 (2020). <https://csrc.nist.gov/publications/detail/nistir/8309/final>.
- [10] Gorjan Alagic, Daniel Apon, David Cooper, Quynh Dang, Thanh Dang, John Kelsey, Jacob Lichtinger, Yi-Kai Liu, Carl Miller, Dustin Moody, Rene Peralta, Ray Perlner, Angela Robinson, Daniel Smith-Tone, Status report on the third round of the NIST Post-Quantum Cryptography Standardization Process, NISTIR 8413 (2022). <https://csrc.nist.gov/publications/detail/nistir/8413/final>.
- [11] Mark Allman, Sally Floyd, Craig Partridge, Increasing TCP's initial window (1998); see also newer version [12]. <https://www.rfc-editor.org/rfc/rfc2414>.
- [12] Mark Allman, Sally Floyd, Craig Partridge, Increasing TCP's initial window (2002); see also older version [11]. <https://www.rfc-editor.org/rfc/rfc3390>.
- [13] ARM, ARM extends 28nm IP leadership with latest UMC 28HPC POPs (2016). <https://web.archive.org/web/20200922064807/https://www.arm.com/company/news/2016/02/arm-extends-28nm-ip-leadership-withlatest-umc-28hpc-pops>.
- [14] Tomer Ashur, Atul Luykx, An account of the ISO/IEC standardization of the Simon and Speck Block Cipher Families (2018). <https://www.esat.kuleuven.be/cosic/publications/article-2957.pdf>.
- [15] Steve Babbage, Christophe De Cannière, Anne Canteaut, Carlos Cid, Henri Gilbert, Thomas Johansson, Matthew Parker, Bart Preneel, Vincent Rijmen, Matthew Robshaw, The eSTREAM portfolio (2008). <https://www.ecrypt.eu.org/stream/portfolio.pdf>.
- [16] Sonia Belaïd, Tim Güneysu (editors), Smart card research and advanced applications—18th international conference, CARDIS 2019, Prague, Czech Republic, November 11–13, 2019, revised selected papers, Lecture Notes in Computer Science, 11833, Springer, 2019. ISBN 978-3-030-42067-3.
- [17] Daniel J. Bernstein, Cache-timing attacks on AES (2005). <https://cr.yp.to/papers.html#cachetiming>.
- [18] Daniel J. Bernstein, Cryptography for the paranoid, slides (2012). <https://cr.yp.to/talks.html#2012.09.24>.
- [19] Daniel J. Bernstein, CAESAR call for submissions (2014). <https://competitions.cr.yp.to/caesar-call.html>.
- [20] Daniel J. Bernstein, Boring crypto, slides (2015). <https://cr.yp.to/talks.html#2015.10.05>.
- [21] Daniel J. Bernstein, CAESAR use cases (2016). <https://groups.google.com/g/crypto-competitions/c/DLv193SPSDc/m/4CeHPvIoBgAJ>.
- [22] Daniel J. Bernstein (editor), Challenges in authenticated encryption, ECRYPT-CSA D1.1, revision 1.05 (2017). <https://chae.cr.yp.to/chae-20170301.pdf>.
- [23] Daniel J. Bernstein, Comparing proofs of security for lattice-based encryption, Second PQC Standardization Conference (2019). <https://cr.yp.to/papers.html#latticeproofs>.
- [24] Daniel J. Bernstein, Visualizing size-security tradeoffs for lattice-based encryption, Second PQC Standardization Conference (2019). <https://cr.yp.to/papers.html#paretoviz>.
- [25] Daniel J. Bernstein, Post-quantum cryptography (2020). <https://cr.yp.to/talks.html#2020.01.30>.
- [26] Daniel J. Bernstein, A discretization attack (2020). <https://cr.yp.to/papers.html#categories>.
- [27] Daniel J. Bernstein, “Agreeing with main points in 3, 4, 6, 10 in <https://eprint.iacr.org/2021/543>. More objections to 2, 5, 7, 9. Most important dispute is regarding risk management, 1+8. Recent advances in torsion-point attacks have killed a huge part of the SIKE parameter space, far worse than MOV vs ECDLP.”, tweet (2021). <https://archive.today/Hiiyi>.
- [28] Daniel J. Bernstein, Tung Chou, Chitchanok Chuengsatiansup, Andreas Hülsing, Eran Lambooj, Tanja Lange, Ruben Niederhagen, Christine van Vredendaal, How to manipulate curve standards: a white paper for the black hat, in SSR 2015 (2015). <https://bada55.cr.yp.to/>.

- [29] Daniel J. Bernstein, Tanja Lange, Security dangers of the NIST curves (2013). <https://cr.yp.to/talks.html#2013.09.16>.
- [30] Daniel J. Bernstein, Tanja Lange, Failures in NIST's ECC standards (2016). <https://cr.yp.to/papers.html#nistecc>.
- [31] Daniel J. Bernstein, Tanja Lange, The year in post-quantum crypto (2018). <https://www.youtube.com/watch?v=q6vnfyfS51w>.
- [32] Daniel J. Bernstein, Tanja Lange (editors), eBACS: ECRYPT Benchmarking of Cryptographic Systems, accessed 24 April 2023 (2023). <https://bench.cr.yp.to>.
- [33] Daniel J. Bernstein, Tanja Lange, McTiny: fast high-confidence post-quantum key erasure for tiny network servers, in USENIX 2020 [41] (2020), 1731–1748. <https://cr.yp.to/papers.html#mctiny>.
- [34] Daniel J. Bernstein, Tanja Lange, Ruben Niederhagen, Dual EC: a standardized back door, in [119] (2015), 256–281. <https://eprint.iacr.org/2015/767>.
- [35] Karthikeyan Bhargavan, Gaëtan Leurent, On the practical (in-)security of 64-bit block ciphers: collision attacks on HTTP over TLS and OpenVPN, in CCS 2016 [135] (2016), 456–467. https://sweet32.info/SWEET32_CCS16.pdf.
- [36] Eli Biham, How to forge DES-encrypted messages in 228 steps, Technion Computer Science Department Technical Report CS0884 (1996). <https://web.archive.org/web/20191118225105/http://www.cs.technion.ac.il/users/wwwb/cgi-bin/tr-info.cgi?1996/CS/CS0884>.
- [37] Eli Biham, A note on comparing the AES candidates, in Second AES Candidate Conference (1999), 85–92. <https://web.archive.org/web/20181006234503/https://cs.technion.ac.il/~biham/publications.html>.
- [38] Dennis K. Branstad (editor), Computer security and the Data Encryption Standard: proceedings of the conference on computer security and the Data Encryption Standard held at the National Bureau of Standards in Gaithersburg, Maryland on February 15, 1977, NBS Special Publication 500-27, 1977. <https://nvlpubs.nist.gov/nistpubs/Legacy/SP/nbsspecialpublication500-27.pdf>.
- [39] Peter Bright, The NSA's work to make crypto worse and better (2013). <https://arstechnica.com/information-technology/2013/09/the-nsaswork-to-make-crypto-worse-and-better/>.
- [40] Christophe De Cannière, eSTREAM Optimized Code HOWTO (2005). <https://www.ecrypt.eu.org/stream/perf/>.
- [41] Srdjan Capkun, Franziska Roesner (editors), Proceedings of the 29th USENIX Security Symposium, USENIX Association, 2020. ISBN 978-1-939133-17-5.
- [42] Matthew R. Carter, Raghurama R. Velagala, John Pham, Jens-Peter Kaps, eXtended eXternal benchmarking eXtension (XXBX), demo at IEEE Hardware Oriented Security and Trust 2018 (2018). https://web.archive.org/web/20210922081539/http://www.hostsymposium.org/host2018/hwdemo/HOST_2017_hwdemo_23.pdf.
- [43] Wouter Castryck, Thomas Decru, An efficient key recovery attack on SIDH, in Eurocrypt 2023 [71] (2023), 423–447. <https://eprint.iacr.org/2022/975>.
- [44] CDN Planet, Initcwnd settings of major CDN providers (2017). <https://www.cdnplanet.com/blog/initcwnd-settings-major-cdn-providers/>.
- [45] Shu-jen Chang, Ray Perlner, William E. Burr, Meltem Sönmez Turan, John M. Kelsey, Souradyuti Paul, Lawrence E. Bassham, Third-round report of the SHA-3 cryptographic hash algorithm competition, NISTIR 7896 (2012). <https://csrc.nist.gov/publications/detail/nistir/7896/final>.
- [46] Jerry Chu, Nandita Dukkkipati, Yuchung Cheng, Matt Mathis, Increasing TCP's initial window (2013). <https://tools.ietf.org/html/rfc6928>.
- [47] Shaanan Cohney, Andrew Kwong, Shahar Paz, Daniel Genkin, Nadia Heninger, Eyal Ronen, Yuval Yarom, Pseudorandom black swans: cache attacks on CTR_DRBG, in S&P 2020 [4] (2020), 1241–1258. <https://eprint.iacr.org/2019/996>.
- [48] Don Coppersmith, The Data Encryption Standard (DES) and its strength against attacks, IBM Journal of Research and Development 38 (1994), 243–250. <https://simson.net/ref/1994/coppersmith94.pdf>.
- [49] Craig Costello, The case for SIKE: a decade of the supersingular isogeny problem (2021). <https://eprint.iacr.org/2021/543>.
- [50] Ronald Cramer (editor), Advances in cryptology–EUROCRYPT 2005, 24th annual international conference on the theory and applications of cryptographic techniques, Aarhus, Denmark, May 22–26, 2005, proceedings, Lecture Notes in Computer Science, 3494, Springer, 2005. ISBN 3-540-25910-4.

- [51] Paul Crowley, Eric Biggers, Adiantum: length-preserving encryption for entry-level processors, IACR Transactions on Symmetric Cryptology 2018 (2018), 39–61. <https://eprint.iacr.org/2018/720>.
- [52] Paul Crowley, Eric Biggers, Introducing Adiantum: encryption for the next billion users (2019). <https://security.googleblog.com/2019/02/introducing-adiantum-encryption-for.html>.
- [53] Cryptographic Technology Group, NIST cryptographic standards and guidelines development process, NISTIR 7977 (2016). <https://csrc.nist.gov/publications/detail/nistir/7977/final>.
- [54] Joan Daemen, Vincent Rijmen, Resistance against implementation attacks: a comparative study of the AES proposals (1999). <https://web.archive.org/web/20000816072451/http://csrc.nist.gov/encryption/aes/round1/conf2/papers/daemen.pdf>.
- [55] Ruth M. Davis, The Data Encryption Standard in perspective, in [38] (1977), 4–13.
- [56] Thomas Decru, Sabrina Kunzweiler, Efficient computation of $(3n, 3n)$ -isogenies (2023). <https://eprint.iacr.org/2023/376>.
- [57] Whitfield Diffie, Martin E. Hellman, Exhaustive cryptanalysis of the NBS Data Encryption Standard, Computer 10 (1977), 74–84. <https://ee.stanford.edu/hellman/publications/27.pdf>.
- [58] Dumitru-Daniel Dinu, Alex Biryukov, Johann Groszschädl, Dmitry Khovratovich, Yann Le Corre, Léo Perrin, FELICS—fair evaluation of lightweight cryptographic systems, NISTWorkshop on Lightweight Cryptography 2015 (2015). <https://hdl.handle.net/10993/25967>.
- [59] Hans Dobbertin, Antoon Bosselaers, Bart Preneel, RIPEMD-160: a strengthened version of RIPEMD, in FSE 1996 [66] (1996), 71–82. <https://homes.esat.kuleuven.be/bosselaer/ripemd160/pdf/AB-9601/AB-9601.pdf>.
- [60] Luan Cardoso Dos Santos, Johann Groszschädl, Alex Biryukov, FELICS-AEAD: benchmarking of lightweight authenticated encryption algorithms, in CARDIS 2019 [16] (2019), 216–233. <https://hdl.handle.net/10993/41537>.
- [61] Orr Dunkelman, Léo Perrin, Adapting rigidity to symmetric cryptography: towards “unswerving” designs, in SSR 2019 (2019). <https://eprint.iacr.org/2019/1187>.
- [62] ECRYPT, Call for stream cipher primitives (2005). <https://www.ecrypt.eu.org/stream/call/>.
- [63] Electronic Frontier Foundation, Cracking DES: secrets of encryption research, wiretap politics & chip design, O’Reilly, 1998. ISBN 978-1565925205.
- [64] Kris Gaj, Jens-Peter Kaps, Venkata Mirineni, Marcin Rogawski, Ekawat Homsirikamol, Benjamin Y. Brewster, ATHENA—Automated Tool for Hardware EvaluationN: toward fair and comprehensive benchmarking of cryptographic hardware using FPGAs, in FPL 2010 [2] (2010), 414–421. https://cryptography.gmu.edu/athena/papers/GMU_FPL_2010_ATHENA.pdf.
- [65] Brian Gladman, Serpent S boxes as Boolean functions (2000). https://web.archive.org/web/20001118002700/http://www.binternet.com/brian.gladman/cryptography_technology/serpent/index.html.
- [66] Dieter Gollmann (editor), Fast software encryption, third international workshop, Cambridge, UK, February 21–23, 1996, proceedings, Lecture Notes in Computer Science, 1039, Springer, 1996. ISBN 3-540-60865-6.
- [67] Google, Addressing performance regressions (2021). https://web.archive.org/web/20211226054614/https://chromium.googlesource.com/chromium/src/+refs/heads/main/docs/speed/addressing_performance_regressions.md.
- [68] Matthew D. Green, “Flipside: What if NIST/NSA know a weakness in 1/10000000 curves? NIST searches space for curves that *aren’t* vulnerable.”, tweet (2013). <https://archive.today/HyWGr>.
- [69] Matthew D. Green, “Discussion with @hashbreaker from when I was younger and more naive. #nist #ecc”, tweet (2013). <https://archive.is/59Hiz>.
- [70] Shai Halevi (editor), Advances in cryptography—CRYPTO 2009, 29th annual international cryptology conference, Santa Barbara, CA, USA, August 16–20, 2009, proceedings, Lecture notes in Computer Science, 5677, Springer, 2009.
- [71] Carmit Hazay, Martijn Stam (editors), Advances in cryptology—EUROCRYPT 2023—42nd annual international conference on the theory and applications of cryptographic techniques, Lyon, France, April 23–27, 2023, proceedings, part V, 14008, Springer, 2023. ISBN 978-3-031-30588-7.
- [72] Martin Hell, Thomas Johansson, Breaking the F-FCR-H stream cipher in real time, in Asiacypt 2008 [108] (2008), 557–569. <https://www.iacr.org/archive/asiacypt2008/53500563/53500563.pdf>.
- [73] Martin E. Hellman, A cryptanalytic time-memory tradeoff, IEEE Transactions on Information Theory 26 (1980), 401–406. <https://ee.stanford.edu/hellman/publications/36.pdf>.

- [74] Martin E. Hellman, Whitfield Diffie, Paul Baran, Dennis Branstad, Douglas L. Hogan, Arthur J. Levenson, DES (Data Encryption Standard) review at Stanford University (1976). <https://web.archive.org/web/20170420171412/www.toad.com/des-stanford-meeting.html>.
- [75] HTTP Archive, Report: state of the web, accessed 24 April 2023 (2023). <https://httparchive.org/reports/state-of-the-web>.
- [76] Bobby R. Inman, The NSA perspective on telecommunications protection in the nongovernmental sector (1979). <https://cryptome.org/nsa-inman-1979.pdf>.
- [77] Akiko Inoue, Tetsu Iwata, Kazuhiko Minematsu, Bertram Poettering, Cryptanalysis of OCB2: attacks on authenticity and confidentiality, *Journal of Cryptology* 33 (2020), 1871–1913. <https://eprint.iacr.org/2019/311>.
- [78] William Jackson, NSA reveals its secret: No backdoor in encryption standard (2011). <https://web.archive.org/web/20141110051727/https://gcn.com/articles/2011/02/16/rsa-11-nsa--no-des-backdoor.aspx>.
- [79] Van Jacobson, Congestion avoidance and control, *ACM SIGCOMM Computer Communication Review* 18 (1988), 314–329.
- [80] David Jao, Luca De Feo, Towards quantum-resistant cryptosystems from supersingular elliptic curve isogenies, in *PQCrypto 2011* [141] (2011), 19–34. <https://eprint.iacr.org/2011/506>.
- [81] Thomas R. Johnson, American cryptology during the cold war, 1945–1989, book III: retrenchment and reform, 1972–1980, 1998. https://archive.org/details/cold_war_iii-nsa.
- [82] Matthias J. Kannwischer, Joost Rijneveld, Peter Schwabe, Ko Stoffelen, pqm4: testing and benchmarking NIST PQC on ARM Cortex-M4 (2019). <https://eprint.iacr.org/2019/844>.
- [83] Jens-Peter Kaps, William Diehl, Michael Tempelmeier, Farnoud Farahmand, Ekawat Homsirikamol, Kris Gaj, A comprehensive framework for fair and efficient benchmarking of hardware implementations of lightweight cryptography (2019). <https://eprint.iacr.org/2019/1273>.
- [84] Richard F. Kayser, Announcing request for candidate algorithm nominations for a new cryptographic hash algorithm (SHA3) family, *Federal Register* 72 (2007), 62212–62220. <https://www.govinfo.gov/content/pkg/FR-2007-11-02/pdf/E7-21581.pdf>.
- [85] Paul Kinnucan, Data encryption gurus: Tuchman and Meyer, *Cryptologia* 2 (1978), 371–381.
- [86] Donald Knuth, Structured programming with go to statements, *Computing Surveys* 6 (1974), 261–301.
- [87] Samuel Kramer, Announcing development of a Federal Information Processing Standard for Advanced Encryption Standard, *Federal Register* 62 (1996), 93–94. <https://www.govinfo.gov/content/pkg/FR-1997-01-02/pdf/96-32494.pdf>.
- [88] Kris Kwiatkowski, Luke Valenta, The TLS post-quantum experiment (2019). <https://blog.cloudflare.com/the-tls-post-quantum-experiment/>.
- [89] Adam Langley, Maybe skip SHA-3 (2017). <https://www.imperialviolet.org/2017/05/31/skipsha3.html>.
- [90] David P. Leech, Michael W. Chinworth, The economic impacts of NIST’s data encryption standard (DES) program (2001). <https://csrc.nist.gov/publications/detail/white-paper/2001/10/01/the-economic-impacts-of-nist-des-program/final>.
- [91] David P. Leech, Stacey Ferris, John T. Scott, The economic impacts of the advanced encryption standard, 1996–2017 (2018). <https://csrc.nist.gov/publications/detail/white-paper/2018/09/07/economic-impactsof-the-advanced-encryption-standard-1996-2017/final>.
- [92] Gaëtan Leurent, Thomas Peyrin, SHA-1 is a shambles: first chosen-prefix collision on SHA-1 and application to the PGP web of trust, in *USENIX 2020* [5] (2020), 1839–1856. <https://eprint.iacr.org/2020/014>.
- [93] Jean Liénardy, Frédéric Lafitte, A weakness in OCB3 used with short nonces allowing for a break of authenticity and confidentiality (2023). <https://eprint.iacr.org/2023/326>.
- [94] Luciano Maino, Chloe Martindale, Lorenz Panny, Giacomo Pope, Benjamin Wesolowski, A direct key recovery attack on SIDH, in *Eurocrypt 2023* [71] (2023), 448–471.
- [95] Stefan Mangard, François-Xavier Standaert (editors), *Cryptographic hardware and embedded systems, CHES 2010, 12th international workshop, Santa Barbara, CA, USA, August 17–20, 2010, proceedings, Lecture Notes in Computer Science*, 6225, Springer, 2010. ISBN 978-3-642-15030-2.
- [96] David McGrew, Impossible plaintext cryptanalysis and probable-plaintext collision attacks of 64-bit block cipher modes (2012). <https://eprint.iacr.org/2012/623>.

- [97] Paul Meissner (editor), Report of the workshop on estimation of significant advances in computer technology, NBSIR 76-1189 (1976). <https://nvlpubs.nist.gov/nistpubs/Legacy/IR/nbsir76-1189.pdf>.
- [98] Kaushik Nath, Palash Sarkar, Efficient 4-way vectorizations of the Montgomery ladder (2020). <https://eprint.iacr.org/2020/378>.
- [99] National Institute of Standards and Technology, NIST's efficiency testing for Round1 AES candidates (1999). <https://web.archive.org/web/20000816072005/https://csrc.nist.gov/encryption/aes/round1/conf2/NIST-efficiency-testing.pdf>.
- [100] National Institute of Standards and Technology, Submission requirements and evaluation criteria for the lightweight cryptography standardization process (2018). <https://csrc.nist.gov/CSRC/media/Projects/Lightweight-Cryptography/documents/final-lwc-submission-requirementsaugust2018.pdf>.
- [101] National Security Archive, "Disreputable if not outright illegal": the National Security Agency versus Martin Luther King, Muhammad Ali, Art Buchwald, Frank Church, et al. (2013). <https://nsarchive2.gwu.edu/NSAEBB/NSAEBB441/>.
- [102] James Nechvatal, Elaine Barker, Lawrence Bassham, William Burr, Morris Dworkin, James Foti, Edward Roback, Report on the development of the Advanced Encryption Standard (AES), Journal of Research of the National Institute of Standards and Technology 106 (2001). <https://nvlpubs.nist.gov/nistpubs/jres/106/3/j63nec.pdf>.
- [103] James Nechvatal, Elaine Barker, Donna Dodson, Morris Dworkin, James Foti, Edward Roback, Status report on the first round of the development of the Advanced Encryption Standard, Journal of Research of the National Institute of Standards and Technology 104 (1999). <https://nvlpubs.nist.gov/nistpubs/jres/104/5/j45nec.pdf>.
- [104] OpenDNS Team, OpenDNS adopts DNSCurve (2010). <https://umbrella.cisco.com/blog/opendns-dnscurve>.
- [105] Dag Arne Osvik, Speeding up Serpent, in Third AES Candidate Conference (2000), 317–329. <https://www.iu.uib.no/osvik/pub/aes3.pdf>.
- [106] Elisabeth Paté-Cornell, Louis Anthony Cox Jr., Improving risk management: from lame excuses to principled practice, Risk Analysis 34 (2014), 1228–1239.
- [107] Nicole Perlroth, Jeff Larson, Scott Shane, N.S.A. able to foil basic safeguards of privacy on web (2013). <https://www.nytimes.com/2013/09/06/us/nsafoils-much-internet-encryption.html>.
- [108] Josef Pieprzyk (editor), Advances in cryptology—ASIACRYPT 2008, 14th international conference on the theory and application of cryptology and information security, Melbourne, Australia, December 7–11, 2008. proceedings, Lecture Notes in Computer Science, 5350, Springer, 2008. ISBN 978-3-540-89254-0.
- [109] Bart Preneel, Antoon Bosselaers, Vincent Rijmen, Bart Van Rompay, Louis Granboulan, Jacques Stern, Sean Murphy, Markus Dichtl, Pascale Serf, Eli Biham, Orr Dunkelman, Vladimir Furman, François Koeune, Gilles Piret, Jean-Jacques Quisquater, Lars Knudsen, Håvard Raddum, Comments by the NESSIE project on the AES finalists (2000). https://www.cosic.esat.kuleuven.be/nessie/deliverables/D4_NessieAESInput.pdf.
- [110] Bart Preneel, Bart Van Rompay, Siddika Berna Örs, Alex Biryukov, Louis Granboulan, Emmanuelle Dottax, Markus Dichtl, Marcus Schafheutle, Pascale Serf, Stefan Pyka, Eli Biham, Elad Barkan, Orr Dunkelman, J. Stolin, Matthieu Ciet, Jean-Jacques Quisquater, Francisco Sica, Håvard Raddum, Matthew Parker, Performance of optimized implementations of the NESSIE primitives (2003). <https://www.cosic.esat.kuleuven.be/nessie/deliverables/D21-v2.pdf>.
- [111] Andrew Regenscheid, Ray Perlner, Shu-jen Chang, John Kelsey, Mridul Nandi, Souradyuti Paul, Status report on the first round of the SHA-3 cryptographic hash algorithm competition, NISTIR 7620 (2009). <https://csrc.nist.gov/publications/detail/nistir/7620/final>.
- [112] Ronald L. Rivest, The MD5 message-digest algorithm, RFC 1321 (1992). <https://tools.ietf.org/html/rfc1321>.
- [113] Matthew Robshaw, The eSTREAM project, in [114] (2008), 1–6.
- [114] Matthew Robshaw, Olivier Billet (editors), New stream cipher designs: the eSTREAM finalists, Lecture Notes in Computer Science, 4986, Springer, 2008. ISBN 978-3-540-68350-6.
- [115] Root Server Operators, Events of 2015-11-30 (2015). <https://root-servers.org/media/news/events-of-20151130.txt>.

- [116] Root Server Operators, Threat mitigation for the root server system (2019). https://root-servers.org/media/news/Threat_Mitigation_For_the_Root_Server_System.pdf.
- [117] Root Server Operators, Statement on DNS encryption (2021). https://root-servers.org/media/news/Statement_on_DNS_Encryption.pdf.
- [118] Alexander Rostovtsev, Anton Stolbunov, Public-key cryptosystem based on isogenies (2006). <https://eprint.iacr.org/2006/145>.
- [119] Peter Y. A. Ryan, David Naccache, Jean-Jacques Quisquater (editors), The new codebreakers: essays dedicated to David Kahn on the occasion of his 85th birthday, Lecture Notes in Computer Science, 9100, Springer, 2015. ISBN 978-3-662-49300-7.
- [120] Jim Salter, A Chrome feature is creating enormous load on global root DNS servers (2020). <https://arstechnica.com/gadgets/2020/08/a-chrome-feature-is-creating-enormous-load-on-global-root-dnsservers/>.
- [121] Bruce Schneier, John Kelsey, Doug Whiting, David Wagner, Chris Hall, Niels Ferguson, Performance comparison of the AES submissions, in Second AES Candidate Conference (1999), 15–34. <https://www.schneier.com/academic/paperfiles/paper-aes-performance.pdf>.
- [122] Michael Scott, Re: NIST announces set of Elliptic Curves (1999). https://groups.google.com/forum/message/raw?msg=sci.crypt/mFMukSsORmI/FpbHDQ6hM_MJ.
- [123] Senate Select Committee on Intelligence, Unclassified summary: Involvement of NSA in the development of the Data Encryption Standard (1978). <https://www.intelligence.senate.gov/sites/default/files/publications/95nsa.pdf>.
- [124] Dimitrios Sikeridis, Panos Kampanakis, Michael Devetsikiotis, Post-quantum authentication in TLS 1.3: a performance study, in NDSS 2020 [3] (2020). <https://eprint.iacr.org/2020/071>.
- [125] Marc Stevens, CWI cryptanalyst discovers new cryptographic attack variant in Flame spy malware (2012). <https://www.cwi.nl/news/2012/cwicryptanalyst-discovers-new-cryptographic-attack-variant-in-flamespy-malware>.
- [126] Marc Stevens, Alexander Sotirov, Jacob Appelbaum, Arjen Lenstra, David Molnar, Dag Arne Osvik, Benne de Weger, Short chosen-prefix collisions for MD5 and the creation of a rogue CA certificate, in Crypto 2009 [70] (2009), 55–69. <https://iacr.org/archive/crypto2009/56770054/56770054.pdf>.
- [127] Nassim Nicholas Taleb, Skin in the game: Hidden asymmetries in daily life, Random House, 2018. ISBN 978-0425284629.
- [128] Haixu Tang, Xiaoqian Jiang, Xiaofeng Wang, Shuang Wang, Heidi Sofia, Dov Fox, Kristin Lauter, Bradley Malin, Amalio Telenti, Li Xiong, Lucila Ohno-Machado, Protecting genomic data analytics in the cloud: state of the art and opportunities, BMC Medical Genomics 9 (2016), article 63. <https://bmcmmedgenomics.biomedcentral.com/articles/10.1186/s12920-016-0224-3>.
- [129] Eran Tromer, Dag Arne Osvik, Adi Shamir, Efficient cache attacks on AES, and countermeasures, Journal of Cryptology 23 (2010), 37–71. <https://link.springer.com/article/10.1007/s00145-009-9049-y>.
- [130] Meltem Sönmez Turan, Kerry McKay, Çağdaş Çarýk, Donghoon Chang, Lawrence Bassham, Status report on the first round of the NIST lightweight cryptography standardization process, NISTIR 8268 (2019). <https://csrc.nist.gov/publications/detail/nistir/8268/final>.
- [131] Meltem Sönmez Turan, Kerry McKay, Donghoon Chang, Çağdaş Çarýk, Lawrence Bassham, Jinkeon Kang, John Kelsey, Status report on the second round of the NIST lightweight cryptography standardization process, NISTIR 8369 (2021). <https://csrc.nist.gov/publications/detail/nistir/8369/final>.
- [132] Meltem Sönmez Turan, Ray Perlner, Lawrence Bassham, William Burr, Donghoon Chang, Shu-jen Chang, Morris Dworkin, John Kelsey, Souradyuti Paul, Rene Peralta, Status report on the second round of the SHA-3 cryptographic hash algorithm competition, NISTIR 7764 (2011). <https://csrc.nist.gov/publications/detail/nistir/7764/final>.
- [133] Visiting Committee on Advanced Technology of the National Institute of Standards and Technology, NIST cryptographic standards and guidelines development process (2014). <https://www.nist.gov/system/files/documents/2017/05/09/VCAT-Report-on-NIST-Cryptographic-Standardsand-Guidelines-Process.pdf>.
- [134] Xiaoyun Wang, Hongbo Yu, How to break MD5 and other hash functions, in Eurocrypt 2005 [50] (2005), 19–35. <https://www.iacr.org/cryptodb/archive/2005/EUROCRYPT/2868/2868.pdf>.

- [135] Edgar R. Weippl, Stefan Katzenbeisser, Christopher Kruegel, Andrew C. Myers, Shai Halevi (editors), Proceedings of the 2016 ACM SIGSAC conference on computer and communications security, Vienna, Austria, October 24–28, 2016, ACM, 2016. ISBN 978-1-4503-4139-4.
- [136] Christian Wenzel-Benner, Jens Gräf, XBX: eXternal Benchmarking eXtension for the SUPERCOP crypto benchmarking framework, in CHES 2010 [95] (2010). https://link.springer.com/chapter/10.1007/978-3-642-15031-9_20.
- [137] Christian Wenzel-Benner, Jens Gräf, John Pham, Jens-Peter Kaps, XBX benchmarking results January 2012 (2012). https://csrc.nist.rip/groups/ST/hash/sha-3/Round3/March2012/documents/papers/WENZEL_BENNER_paper.pdf.
- [138] Michael Wertheimer, Encryption and the NSA role in international standards, Notices of the AMS (2015), 165–167. <https://www.ams.org/notices/201502/moti-p165.pdf>.
- [139] Bas Westerbaan, Sizing up post-quantum signatures (2021). <https://web.archive.org/web/20211123165010/https://blog.cloudflare.com/sizing-up-post-quantum-signatures/>.
- [140] Michael J. Wiener, Efficient DES key search, Carleton University School of Computer Science Technical Report TR-244 (1994). <https://carleton.ca/scs/wp-content/uploads/TR-244.pdf>.
- [141] Bo-Yin Yang (editor), Post-quantum cryptography: 4th international workshop, PQCrypto 2011, Taipei, Taiwan, November 29–December 2, 2011, proceedings, 7071, Springer, 2011. ISBN 978-3-642-25404-8.
- [142] John Young, NSA FOIA documents on Joseph Meyer IEEE letter (2010). <https://cryptome.org/0001/nsa-meyer.htm>.
- [143] Kim Zetter, How a crypto ‘backdoor’ pitted the tech world against the NSA (2013). <https://www.wired.com/2013/09/nsa-backdoor/>.

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.