From: D. J. Bernstein <djb@cr.yp.to>
Sent: Monday, September 21, 2020 5:50 AM
To: pqc-comments
Cc: pqc-forum
Subject: ROUND 2 OFFICIAL COMMENT: NTRU Prime
Attachments: signature.asc

I'm writing on behalf of the NTRU Prime team, as a public response to NIST's request for a summary of expected changes in round 3.

NTRU Prime is a small lattice system. Subject to this constraint, our primary objective is to eliminate unnecessary complications in security review. We correctly predicted that such complications would lead to security failures in NISTPQC lattice submissions. We evaluated a variety of trapdoor functions from this perspective before submission, again during round 1, and again during round 2.

On this basis we have once again decided against decryption failures; modules; errors; and all other changes that we have considered to our family of trapdoor functions. We therefore plan to submit the same family of trapdoor functions in round 3. NTRU Prime will therefore have an unchanged family of trapdoor functions throughout round 1, round 2, and round 3.

Our CCA conversion includes various hashing safeguards, some already in round 1 and some added in round 2. These safeguards cost 32 bytes in ciphertext size and a considerable fraction of our CPU time. However, even with these safeguards, NTRU Prime often outperforms other small lattice KEMs, as the following references show:

https://cr.yp.to/papers.html#paretoviz
https://cr.yp.to/talks/2019.08.23-1/slides-djb-20190823-1-paretoviz-4x3.pdf
https://bench.cr.yp.to/results-kem.html#amd64-hiphop
https://github.com/mupq/pqm4/blob/master/benchmarks.md

More importantly, the costs of our hashing safeguards are negligible in applications. We plan to submit the same CCA conversion in round 3.
NTRU Prime will thus be fully compatible between round 2 and round 3, when users choose the same parameters.

Regarding parameter selection, we are concerned that pre-quantum Core-SVP levels 2^100, 2^106, and 2^111, proposed for category 1 for Dilithium, NTRU, and Kyber respectively, will turn out to be inadequate against generic lattice attacks. We will not add dimensions below our
653 (pre-quantum Core-SVP 2^129). We recommend our original dimension
761 (pre-quantum Core-SVP 2^153) for an extra security margin.

We have seen various requests for larger dimensions, even larger than our dimension 857 (pre-quantum Core-SVP 2^175). To accommodate these requests and prevent any accusations of a lack of flexibility, we plan to add some larger dimensions as a supplement to our current dimensions.

We have also considered adding intermediate parameter sets to further illustrate our flexibility, showing that NTRU Prime offers even larger advantages in Core-SVP under various size limits compared to, e.g., Kyber. The call for proposals explicitly allowed multiple parameter sets per category. However, NIST has now made an announcement that seems to discourage "too many parameter sets".

The rest of this message is regarding the problems caused by NIST's unstable definitions of security categories. These are problems for the NISTPQC process broadly, not just for NTRU Prime.

The call for proposals specified AES-128 key search as a "floor" for category 1 in "all metrics that NIST deems to be potentially relevant to practical security". The call for proposals similarly specified floors for other categories. However, this is not a clear and stable category definition unless the metrics are clear and stable.

As an illustration of how much impact metrics have, there is a 40-year literature studying metrics for realistic large-scale two-dimensional models of computation. Standard theorems---see, e.g.,

https://www.eecs.harvard.edu/~htk/publication/1981-jacm-brent-kung.pdf

---imply that these metrics assign 50% higher asymptotic exponents to large-integer multiplication, large-array sorting, etc. than a "gates"
metric does. (Analogous three-dimensional metrics studied in, e.g.,

https://link.springer.com/article/10.1007/BF01744565

are for machines that appear far more difficult to build than quantum computers, and still have 33% higher exponents than a "gates" metric.) Any attack that has large-scale sorting as a bottleneck is affected by this, whereas AES-128 key search is not.

The call for proposals highlighted "classical gates" and "quantum gates"
(with limited depth) as metrics. However, NIST is not requiring lattice submissions to meet the "classical gates" floor. (See examples below.)

NIST also has not defined a replacement metric for submissions to use.
All lattice submissions have Core-SVP evaluations, but AES-128 does not.
Core-SVP is not a metric for the cost of computation: it is a mechanism for claiming security levels (in an undefined metric) specifically for lattices. Ray Perlner's message dated 9 Jun 2020 15:39:09 +0000 stated "we feel that the CoreSVP metric does indicate which lattice schemes are being more and less aggressive in setting their parameters", but the mapping from Core-SVP evaluations to categories remains undefined.

NIST IR 8309's handling of categories is not consistent across lattice submissions. Consider the following examples from three different
submissions:

(P80) Category 3 for pre-quantum Core-SVP 2^153; 153 is 80% of 192.

(P78) Category 1 for pre-quantum Core-SVP 2^100; 100 is 78% of 128.

(P71) Category 3 for pre-quantum Core-SVP 2^136; 136 is 71% of 192.

P78 is the lowest of these three examples in Core-SVP, and P71 is the lowest in Core-SVP as a percentage of the AES key size for the category.
However, NIST's wording was strikingly more negative for P80 than for
P78 and P71:

(P80) "quite aggressive compared to most of the other submissions
targeting the same security categories"; need to study "whether
they actually meet their claimed security categories";

(P78) "lowest CoreSVP security strength parameter set of any of the lattice schemes still in the process"; need more study on "understanding the concrete security";

(P71) "lower CoreSVP complexity than many of the other schemes targeting the same security strength categories"; need to "understand exactly ... bit security strengths".

Notice, e.g., that NIST asks whether P80 "actually" meets its "claimed" security category, while NIST does not ask the same question regarding P78 or P71.

If NIST were applying the "classical gates" metric then none of P80, P78, and P71 would be able to confidently claim these categories. For example, the uncertainties in Core-SVP seem very unlikely to turn Core-SVP 2^100 into 2^143 "classical gates". Most of the remaining lattice submissions (at least Dilithium, NTRU, Kyber, and NTRU Prime; perhaps also SABER after the announcement that SABER's security levels were miscalculated) would have to adjust their category assignments.
Even worse, some of these submissions (at least Dilithium, NTRU, and Kyber) would have to remove some previously proposed parameters, which seems contrary to the idea of being ready for standardization.

All of these submissions argue, with varying levels of detail and references, that the "classical gates" metric underestimates the actual cost of known attacks. NIST seems receptive to the idea of using a more realistic metric, but has taken four years to post its "preliminary thoughts" on the realism of several different metrics. It is not clear what metrics NIST will end up defining, and it is not clear how long NIST will take to settle on the definitions. What is clear is that NIST has not applied the categories consistently, as illustrated by NIST IR 8309 assigning P80 more negative wording than P78 and P71.

The different wording regarding P80, P78, and P71 appears to have translated into different action, and this seems particularly important for NIST's handling of NTRU Prime. As context, NIST IR 8309 describes finalists in general as

> the most promising to fit the majority of use cases and most likely to be ready for standardization soon after the end of the third round.

We have shown that NTRU Prime fits practically all use cases. As far as we can tell, beyond general concerns about the safety of lattice-based cryptography and about the safety of all small lattice proposals, NTRU Prime is ready for standardization now with our existing parameter sets.
The only negative comments that NIST IR 8309 made regarding NTRU Prime were regarding parameter sets. Specifically, NIST seemed to criticize

* NTRU Prime's assignment of pre-quantum Core-SVP 2^153 to Category 3 (this is exactly P80 above),

* NTRU Prime's assignment of _post-quantum_ Core-SVP 2^159 (pre-quantum Core-SVP 2^175) to Category 4 (this has a larger security margin than P80),

* NTRU Prime's assignment of _post-quantum_ Core-SVP 2^117 (pre-quantum Core-SVP 2^129) to Category 2 (this has a larger security margin than P80), and

* NTRU Prime's "narrower range of CoreSVP values" (our understanding now is that this wasn't a negative comment but merely a request for larger parameters going forward).

Meanwhile various lattice submissions with objectively more dangerous parameter selections were given less critical wording by NIST and were selected as finalists. We see no explanation for why NIST treated P78 and P71 in those submissions more gently than P80 in NTRU Prime.

An application limited to 1024 bytes for keys and plaintexts reaches Core-SVP 2^129 with NTRU Prime's proposed parameters and nothing better than 2^111 with Kyber's proposed parameters. 2^129 is higher security relative to category 2 than 2^111 relative to category 1, and obviously higher security on an absolute scale. NIST's report did not acknowledge this security advantage of NTRU Prime.

We are concerned that the lack of clear, stable, consistently applied category definitions will be used in the continuation of NISTPQC to again make NTRU Prime's parameter choices artificially sound worse than more dangerous parameter choices in other submissions. If we try to reduce this risk by downgrading (e.g.) 2^129 to category 1, while Kyber is allowed to remain in category 1 with just 2^111, then NTRU Prime will be unfairly punished in performance comparisons.

We request that NIST issue clear and stable definitions of the metrics used to define NIST's security categories. At this point in the NISTPQC process, clarity and stability are more important than the exact level of realism. Beyond the floor for the categories, one can reasonably argue that users should take higher Core-SVP levels for all lattice submissions in light of continued advances in lattice attacks; but NIST should handle this in a way that is fair to all submissions. As soon as the evaluation criteria are made clear, we will be happy to adjust our category assignments accordingly.

---Dan

From: D. J. Bernstein <djb@cr.yp.to>
Sent: Sunday, September 27, 2020 4:43 AM
To: pqc-forum
Cc: pqc-comments
Subject: ROUND 2 OFFICIAL COMMENT: NTRU Prime
Attachments: signature.asc

This message has three questions for NIST.

In https://www.youtube.com/watch?v=CBGX1OMzN1o&t=37m55s a few days ago, NIST stated "We're we're still uh have some some questions about NTRU Prime" but didn't elaborate. What are NIST's questions about NTRU Prime?

The late notification and lack of information are problematic. NIST has asked for round-3 tweaks by 1 October, which is just a few days from now. Did I miss some NIST publication listing NIST's questions?

I see only one part of NIST IR 8309 that can be understood as a question about NTRU Prime: namely, "whether they actually meet their claimed security categories will need to be determined" regarding the parameter choices. Meanwhile NIST did not ask the same question regarding other submissions that have objectively more dangerous parameter selections.
The NTRU Prime team email dated 21 Sep 2020 11:49:53 +0200 gave examples of this phenomenon. Procedurally, seeing the issue raised by NIST was a prerequisite for responding to it (and the difficulty of responding was exacerbated by the lack of clarity regarding NIST's "categories").

This section of the talk appeared to be presenting NIST's rationale for selecting lattice finalists and lattice alternates. It would thus seem that the existence of NIST's "questions" regarding NTRU Prime played a role in this. Why didn't NIST IR 8309 say this and provide the list of questions?

---Dan (speaking for myself)