How strong are strong primes? A proposal of test

Raúl DURÁN DÍAZ
Tratamiento de la Información y Codificación, Instituto de Física Aplicada, C.S.I.C.
C/ Serrano 144, E-28006 Madrid, Spain

and

Jaime MUÑOZ MASQUÉ
Tratamiento de la Información y Codificación, Instituto de Física Aplicada, C.S.I.C.
C/ Serrano 144, E-28006 Madrid, Spain

ABSTRACT

After revisiting a number of Internet protocols whose security rests on the RSA cryptosystem, we bring about the discussion on the necessity of using special primes for it to work at the best possible level of security. We propose and discuss a property of integer numbers that permits the characterization of a set of special primes particularly well suited for the RSA cryptosystem.

Keywords: Strong Primes Optimality, Strong Primes, Gordon algorithm, RSA Cryptosystem, TCP/IP Protocols Security.

1. INTRODUCTION

As the XXI century strides along, society has assigned information a particularly important role. In many cases, information has become the most valuable asset for corporations, organizations and even for individuals.

At the same time we attend a most prodigious revolution in the telecommunications world, driven by developments like the world wide Internet that, linking thousands of millions of computers, makes it possible to deliver information stored virtually anywhere, no matter how far, just at any user’s mouse click.

The multifarious possibilities opened by a widespread use of Internet have put pressure on the issue of warranting a controlled access to information and have boosted the importance of secure and private communications over the world wide network.

The ubiquitous HTTP protocol has fueled the development of its secure companion S-HTTP, both stacked on top of TCP/IP, working at application level. The latter was soon taken over by other enhanced solutions like SSL (developed by Netscape), and its competitor PCT (developed by Microsoft), that work at the transport level, thus making encryption a transparent process for any application layered on top of them. More recently, the efforts devoted to develop an IP level encryption protocol, yet one step down TCP/IP stack, have yielded the IPsec specification, implemented in protocols like S/WAN.

In the world of electronic commerce, secure and private communications are an obvious must and so several proposals have seen the light as of late: iKP by IBM, SET by VISA and MasterCard and Secure Courier, by Netscape.

Almost all the protocols cited above are grounded on a small number of cryptographic systems: DES, for secret key techniques and RSA for public key schemes. This fact stresses the importance of assessing their reliability and strength against the attacks of cryptanalysts and hackers if we are willing to convert the expectations of e-commerce, internet banking, etc. into really profitable activities.

In particular, we want to focus our attention on the RSA cryptosystem, since it is used today almost universally to carry out the negotiation of a secret session key between a client and a server. We will discuss in deeper detail the relevance of using special primes—the so-called “strong primes”—and propose a property of positive integers that could help to gauge the, so to speak, “level of strength” of these special primes.

2. STRONG PRIMES

The most common definition of a strong prime is the following:

Definition 1: A prime number $p$ is said to be “strong” (cf. [4, 4.52])—sometimes called “Gordon secure primes” (cf. [2], [5])—if:

---

1 Supported by CICYT (Spain) under grant TE95–1020.
(1) $p - 1$ has a large prime factor $r$, say $p - 1 = ra$,
(2) $p + 1$ has a large prime factor $s$, say $p + 1 = sb$,
(3) and $r - 1$ has a large prime factor $t$, say $r - 1 = tc$.

From the seminal paper [7] of the RSA cryptosystem, the use of strong primes for the factors of the RSA modulus $n = p \cdot q$ was initially encouraged in order to protect the system against powerful algorithms such as [6], [10], generically known as $p \pm 1$ methods of factorization.

However, the discussion regarding the necessity or utility of using this kind of special primes remains open. Some authors—remarkably [8]—argue that “strength” is not necessary. Their arguments can be summarized into two classes:

(i) Using prime numbers satisfying the conditions (1)-(3) adds very little to the system security since, while protecting against the above mentioned $p \pm 1$ methods of factorization ([6], [10]) they do not protect against its generalization: Lenstra’s algorithm of the elliptic curve [3];

(ii) the cyclic attacks are essentially a factoring attack, but less effective, since they are not used as a general purpose factoring method. Moreover, the chance that a cycling attack will succeed is negligible, even if the primes are chosen randomly.

As for point (i), it is noticeable that, in [8], no estimate of the probability that conditions (1)-(3) be fulfilled in a purely random selection, is provided. Besides, one must be sure that $p$ and $q$ do not belong to the classes of prime numbers for which Pollard’s and Williams’s algorithms are very efficient. Finally, the elliptic curve algorithm does not work efficiently for the current lengths of the prime factors of the modulus. However, in [8] the authors do provide a table reporting the probabilities of success of both $p \pm 1$ and elliptic curve methods and conclude stating that “1024-bit RSA moduli are well beyond the computing range of the $p \pm 1$ and elliptic curve methods within the lifetime of the universe”.

As for point (ii), the arguments in [8] may be nullified if $p$ or $q$ are carelessly chosen because if all prime factors of $r - 1$ (see Definition 1) happen to be small, then the cyclic attack becomes efficient. This is, however, extremely unlikely according to [1], a reference where the authors prove that the likelihood of success for a cycling attack is asymptotically nil. Actually, they show that, if we set the modulus $N = p \cdot q$, and we choose $e$ as the public exponent, and $M$ as a message, then for almost all choices of $p, q, e$, and $M$, the smallest cycling value $k$ (i.e. the value of $k$ such that $M = C^{k-1} \mod N$) is nearly as large as $N$ itself.

Still, all these arguments are probabilistic in nature: they do not guarantee that a particular choice of the system parameters will yield a cyclic-attack-proof system. The question remains open of assessing the value of the cycling value $k$ in terms of the chosen parameters, to avoid those that may compromise the security of the system.

Other authors (e.g. [4]), though sharing the idea that there is no definite reason for using strong primes, still recommend them since they offer an enhanced security at almost no extra cost. Actually, Gordon estimates in [2] that finding a strong prime is only 19% harder in time than finding a mere random prime.

3. OPTIMALITY OF A STRONG PRIME

We introduce now a property of integers allowing us to characterize the “level of strength” of a prime. We say that a strong prime is “optimal” if the integers $r$, $s$, and $t$ in Definition 1 above are as large as possible. Hence a natural way to measure the “optimalitjy” of a strong prime is to check how small is the value

$$
\sigma = a + b + c = \frac{p - 1}{r} + \frac{p + 1}{s} + \frac{r - 1}{t}
$$

In fact, we have that the following two propositions hold:

Proposition 1: If $p$ admits a decomposition such that the primes $r$, $s$, and $t$ are odd, then $\sigma \geq 12$.

Proof: Remark that $a, b, c$ are even. The case $\sigma = 6$ is readily seen to be impossible. Then, the cases $\sigma = 8$ and $\sigma = 10$ are discarded after a not difficult but long case-by-case discussion, assigning the values 2, 4, or 6 to $a, b, c$ in all the possible ways. □

The remaining cases are covered by the following

Proposition 2: If $p$ does not satisfy the hypothesis of the Proposition 1, then either $p$ is a Fermat prime or $p$ is a Mersenne prime, or all of the odd prime factors of $p - 1$ are Fermat primes.

Proof: It is not difficult to see that, otherwise, $p$ would admit a decomposition such as the one required to satisfy the hypothesis in the Proposition 1. □

For a given prime $p$ satisfying (1)-(3), the sum $a + b + c$ takes the lowest value when $r$, $s$, $t$ are chosen to be the largest prime factors of $p - 1, p + 1, r - 1$, respectively. Accordingly, let $S(n)$ be the largest prime factor of an integer $n$, usually called the “smoothness” of $n$ (e.g., see [9]). We define a function $\sigma$: $N \setminus \{1, 2\} \rightarrow \mathbb{N}$ as

$$
\sigma(n) = \frac{n - 1}{S(n - 1)} + \frac{n + 1}{S(n + 1)} + \frac{S(n - 1) - 1}{S(n - 1) - 1}
$$

Equipped with these tools, we state the following
Theorem: For every prime number \( p \geq 23 \), \( \sigma(p) \geq 12 \). Hence a strong prime is optimal if and only if \( \sigma(p) = 12 \).

Proof: If \( p \) satisfies the assumptions in Proposition 1, then the statement follows. If \( p \) does not satisfy these assumptions, then from Proposition 2 we are led to distinguish the following cases:

1. If \( p \) is a Fermat prime, then \( p = 1 + 2^a \), so 
   \[ S(p - 1) = 2; \]
   \[ \frac{p - 1}{S(p - 1)} = 2^{a - 1} \]
   and by virtue of the hypothesis,
   \[ \sigma(p) \geq 2^{a - 1} + 2 = \frac{p - 1}{2} + 2 \geq 12. \]
   where we have taken into account that \( a, b, c \) are at least 1, and so the sum of two of them must be at least 2. Since \( p \geq 23 \), the statement follows.

2. If \( p \) is a Mersenne prime, then \( p = 2^b - 1 \), so 
   \[ S(p + 1) = 2; \]
   \[ \frac{p + 1}{S(p + 1)} = 2^{b - 1} \]
   and
   \[ \sigma(p) \geq 2^{b - 1} + 2 = \frac{p + 1}{2} + 2 \geq 12. \]
   Again, the statement follows since \( p \geq 23 \).

3. Last, if \( p - 1 \) is such that all of its odd prime factors are Fermat primes, then necessarily \( S(p - 1) = 1 + 2^c \); therefore, \( S(S(p - 1) - 1) = 2 \), and
   \[ S(p - 1) - 1 \]
   \[ S(S(p - 1) - 1) = 2^{c - 1}. \]
   As the smallest Fermat prime is 3, it is obvious that \( (p + 1)/S(p + 1) \geq 2 \), and hence
   \[ \sigma(p) \geq 1 + \frac{p - 1}{1 + 2^c} + 2 + 2^{c - 1} \geq 12. \]
   It can be readily seen that this statement holds for any \( c \geq 0 \) provided that \( p \geq 41 \). Using the following table, where the values of \( \sigma(p) \) are computed for the primes 23 \( \leq p \leq 41 \):

\[
\begin{array}{c|c|c|c|c|c}
\hline
p & 23 & 29 & 31 & 37 & \sigma(p) \\
\hline
\end{array}
\]

we reach the desired conclusion.

The proof is then finished for all possible cases. \( \Box \)

Proposition 3: For any prime \( p > 3 \) it is true that 24 divides \( p^2 - 1 \).

Proof: Since \( p \) is prime, it is odd, so \( p = 2m + 1 \).

Hence, \( p^2 - 1 = 4m(m + 1) \); moreover either \( m \) or \( m + 1 \) will be even, and therefore, \( p^2 - 1 \) is a multiple of 8.

Let now \( m \equiv 1 \mod 3 \); \( \text{i.e., } m = 1 + 3l \). By substituting we obtain

\[ p = 2m + 1 = 2(1 + 3l) + 1 = 3(1 + 2l), \]

but this is impossible by virtue of the hypothesis. Therefore, either \( m \equiv 0 \mod 3 \) or \( m \equiv 2 \mod 3 \). In both cases, the product \( m(m + 1) \) is a multiple of 3, thus concluding. \( \Box \)

Taking this result into account, we finally obtain the desired characterization:

Corollary: A prime \( p > 29 \) is an optimal strong prime if and only if the following conditions hold:

1. \( (p - 1)/6 \) is 1-safe,
2. \( S(p - 1) = (p - 1)/6 \), and
3. \( S(p + 1) = (p + 1)/4 \).

Proof: A prime is strong if all the conditions in Definition 1 are verified. Since \( p \geq 23 \), we may safely assume that it is always possible to choose \( r \) and \( s \) so that \( r > 3 \) and \( s > 3 \). To fix ideas, if we suppose that \( r \) can be chosen only as 3 or as 2, then this means that either \( p - 1 - 3^a \) or \( p - 1 - 2^b \), \( a \geq 3 \), \( b \geq 5 \). A boundary for \( \sigma(p) \) for each case will then be:

\[ \sigma(p) \geq \frac{p - 1}{S(p - 1)} + 4 = 3^{a - 1} + 4 \geq 3^2 + 4 \geq 13 \]

or

\[ \sigma(p) \geq \frac{p - 1}{S(p - 1)} + 4 = 2^{b - 1} + 4 \geq 2^3 + 4 \geq 20. \]

But then it becomes apparent that a prime \( p \) with such a factorization can never be optimal. The same conclusion can be achieved if we work with the \( s \), supposing now that either \( p + 1 = 3^a \) or \( p + 1 = 2^b \). Assume, then, that \( r > 3 \) and \( s > 3 \). Now, from Proposition 3, we know that \( p^2 - 1 \) is a multiple of 24, and \( p^2 - 1 = 24k = (p + 1)(p - 1) = rsa_b \), so it must be \( ab = 24 \), since \( r > 3 \), \( s > 3 \). As all \( a, b, c \) are even, and \( \sigma(p) = a + b + c = 12 \), then either \( a = 4, b = 6, c = 2 \) or \( a = 6, b = 4, c = 2 \). But the first choice is not possible since, then, we would have:

\[ p = 1 + 4r = 1 + 4(1 + 2t) = 5 + 8t = -1 + 6s, \]

so \( 8t = 6(s - 1) \). If \( t = 3 \), then \( s = 5 \) and hence \( p = -1 + 6s = 29 \), which is impossible; therefore \( t > 3 \) and this is a contradiction for \( r \) is 1-safe. \( \Box \)
4. CONCLUDING REMARKS

The lower bound $\sigma(p) \geq 12$ is specific of prime integers. In fact, there exist many non-prime integers $n$ such that $\sigma(n) < 12$; for example, if $p$ is a 2-safe prime (i.e., $p = 2p_1 + 1, p_1 = 2p_2 + 1, p_1, p_2$ being prime), then $n = 2p - 1$ satisfies $\sigma(n) = 8$.

It is remarkable that a major part of the points of the curve in the graph $n \mapsto \sigma(n)$, $3 \leq n \leq 1200$, remain under 100. In fact, the percentage of values of $n$ such that $\sigma(n) \leq 100$, is 86.25%. In spite of this, peaks in this curve—the numbers opposite to be optimal—keep appearing.

Similar comments apply to the graph of $p \mapsto \sigma(p)$, $3 \leq p \leq 1200$. In this case, the percentage of points under 100 is 78.97%.

For every real number $x \geq 0$, let $\pi_\sigma(x)$ be the number of optimal prime numbers $p \leq x$. Then, we obtain a counting function $\pi_\sigma : [0, +\infty) \rightarrow \mathbb{N}$. While the graph of $\pi_\sigma$ displays that the density of optimal strong primes is rather low, the curve maintains an increasing profile thus making us think that there exist infinitely many optimal prime numbers.

As said before, in [2], Gordon proposes a simple technique that supplies strong primes easily. However, Gordon gives no indication about the “level of strength” of the primes thus computed. The $\sigma$ function becomes the natural candidate for measuring such “level of strength”.

Following Gordon’s technique, the next table has been generated:

<table>
<thead>
<tr>
<th>$p$</th>
<th>20317457717</th>
<th>1146917693</th>
<th>1098831893</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sigma(p)$</td>
<td>773268</td>
<td>336670</td>
<td>234952</td>
</tr>
<tr>
<td>$p$</td>
<td>421265123</td>
<td>2354812133</td>
<td>1335796561</td>
</tr>
<tr>
<td>$\sigma(p)$</td>
<td>216106</td>
<td>34914</td>
<td>607662</td>
</tr>
<tr>
<td>$p$</td>
<td>1521089387</td>
<td>1536206897</td>
<td>1350012737</td>
</tr>
<tr>
<td>$\sigma(p)$</td>
<td>17099710</td>
<td>72442</td>
<td>1400554</td>
</tr>
<tr>
<td>$p$</td>
<td>17849490979</td>
<td>183006137</td>
<td>15712927081</td>
</tr>
<tr>
<td>$\sigma(p)$</td>
<td>408856</td>
<td>64708</td>
<td>1455392</td>
</tr>
</tbody>
</table>

The computed values of $\sigma(p)$ for each $p$ let clearly see that Gordon’s technique, though good in view of computation time, yields poor strong primes measured in terms of the $\sigma$ function: remark that the computed $\sigma$ is at least 34914, very far from the minimum 12 proved above. This leaves open the field to the research of new methods and techniques for the generation of strong prime displaying a good “level of strength”.

5. REFERENCES


