COMPUTATIONAL ASPECTS IN THE GENERATION OF HIGHER-ORDER SAFE PRIMES

R. Durán Díaz¹, L. Hernández Encinas², J. Muñoz Masqué²

¹Departamento de Automática, Universidad de Alcalá
²Instituto de Física Aplicada, CSIC

e-mail(s): raul.duran@uah.es, luis@iec.csic.es, jaime@iec.csic.es

Spain

Abstract: First, an introduction on the current trends of research about special primes is provided. Then, the definition and basic properties of safe primes are presented, extending the concept to higher-order safe primes. An explicit formula to compute the density of this class of primes in the set of the integers is also presented. Finally, explicit conditions are provided permitting the computation of safe primes of arbitrary order.

Key words: Distribution of safe primes; safe-prime signature; chains of primes; Hardy-Littlewood and Bateman-Horn conjectures; Public Key Cryptography.

1. INTRODUCTION

Several popular public-key cryptosystems base their security on the problem of factorising a (typically) two-factor number. A well-known example is RSA, which has been extensively analysed in this regard (see, for example, (Boneh, 1999)) and for which the equivalence between breaking the cryptosystem and factorising the modulus has been proved to be definitely not true when the private exponent is low (see (Boneh and Durfee, 1999), (Boneh and Venkatesan, 1998)). In fact, it has been recently proved that the knowledge of RSA \((e,d)\) parameters yields the factorisation of the modulus \(n\) in deterministic polynomial time (see (Coron and May, 2007)), a result much longed for.

¹ Partially supported by CDTI, Ministerio de Industria, Turismo y Comercio (Spain) in collaboration with Telefónica I+D, Project SEGUR@ with reference CENIT-2007 2004 and Ministerio de Educación Ciencia of Spain under MTM2005-00173.
At the same time, new factorisation methods have been searched for and tried, which, if proved effective, could be a severe drawback for the practical use of this type of cryptosystems.

The factorisation algorithms can be broadly divided into two groups: General Number Field Sieve, Quadratic Sieve, Elliptic Curve, etc., and the so-called $p \pm 1$ algorithms, which try to factor $p+1$ or $p-1$ for some factor $p$ of the number $n$ under test. Among this type of algorithms, Pollard’s (Pollard, 1974) and Williams’ (Williams, 1982) algorithms are the most remarkable ones. While no precise analysis exists about their running times, the $p \pm 1$ algorithms are particularly useful when all the prime factors of $p+1$ are small, and all the prime factors of $p+1$ are small but for one factor $q$, such that $q+1$ or $q-1$ has small prime factors.

Primes not fulfilling the conditions above might be called “strong”, since they are resilient to this kind of factorisation attack, but authors fail to agree on a precise definition. Closely related to the notion of “strong” prime, there exists also the notion of “safe” prime, for which a lot of literature can be found as well: see, for example, (Cai, 1994), (Camenisch and Michels, 1999), (Naccache, 2003), (Wiener, 2003). The so-called Sophie Germain primes, introduced by that author, are in direct connection with safe primes and have been studied since its inception, back in the XIX century (see, for example, (Dubner, 1996), (Yates, 1991)).

There has been a lot of discussion about the convenience of using safe primes for certain applications: It has been recommended, for example, for the modulus of RSA (Menezes et al., 1997), in order to avoid the attacks of the $p+1$ algorithms. Another possible application is the BBS (Blum et al., 1986). It has been proved that to obtain the longest cycle lengths for this generator, it suffices to have $n = p \cdot q$, where both $p$ and $q$ are safe primes (more precisely, 2-safe primes, whose definition will be introduced later on) and $n$ is the modulus for BBS. With additional conditions, it is also possible to use 1-safe primes (to be also defined later on) and still obtain maximal cycle lengths (see (Hernández Encinas et al., 1997)).

As a natural extension, we have introduced also the concept of higher-order safe primes, for which the concept of signature is applicable. In fact, only the positive signature has received attention in the literature (see, for example, (Maurer, 1990), (Maurer, 1992), (Maurer, 1995), (Mihailescu, 1994)). For this reason, we have turned toward a deeper research on the case of negative and mixed signatures.

2. SAFE PRIMES

In this Section we introduce the definition of general safe primes and some elementary property. We will skip all the proofs, since they are rather lengthy.
2.1. Definition and elementary properties

Definition 1 An odd prime integer \( p \) is said to be \( k \)-safe with signature \( \varepsilon_1,...,\varepsilon_k \), with \( \varepsilon_1,...,\varepsilon_k \in \{+1,-1\} \), if \( k \) odd prime integers \( p_1,...,p_k \) exist such that \( p = 2p_1 + \varepsilon_1, p_2 + \varepsilon_2,..., p_{k-1} = 2p_k + \varepsilon_k \). The integer \( k \) is termed the order of the signature.

Notation 2 The set of prime numbers is denoted by \( \mathbf{P} \). For every \( k > 0 \), the set of \( k \)-safe prime integers with signature \( \varepsilon_1,...,\varepsilon_k \) is denoted by \( \mathbf{P}(\varepsilon_1,...,\varepsilon_k) \). We set \( \mathbf{P}^+_k = \mathbf{P}(\varepsilon_1,...,\varepsilon_k) \), when \( \varepsilon_1 = ... = \varepsilon_k = +1 \), and \( \mathbf{P}^-_k = \mathbf{P}(\varepsilon_1,...,\varepsilon_k) \), when \( \varepsilon_1 = ... = \varepsilon_k = -1 \).

Proposition 3 If \( p > 5 \) is a \( k \)-safe prime integer with signature \( \varepsilon_1,...,\varepsilon_k \), then

\[
p = 2^k + \sum_{i=1}^{k} \varepsilon_i 2^{k-i} \quad (\text{mod} \ 2^{i+1})
\]

The proof is essentially carried out proceeding by induction on \( k \).

Corollary 4 If \( \mathbf{P}(\varepsilon_1,...,\varepsilon_k) = \emptyset \), then \( \mathbf{P}(\varepsilon_l,...,\varepsilon_j) = \emptyset \) for \( l > k \).

Example 5 In the literature, the term 1-safe prime refers normally to the set defined by us as \( \mathbf{P}^+_1 \), having signature +1. We have also 1-safe primes with signature −1, namely, the primes in the set \( \mathbf{P}^-_1 \). The former are congruent to 3 modulo 4, whereas the latter are congruent to 1 modulo 4.

Remark 6 The order of the components in the signature is relevant, because it can occur that \( \mathbf{P}(\varepsilon_1,...,\varepsilon_k) \neq \mathbf{P}(\varepsilon_{\pi(1)},...,\varepsilon_{\pi(k)}) \) for a given permutation \( \pi \) of \( \{1,...,k\} \). For example, \( 7 \in \mathbf{P}(+1,-1) \) and \( 7 \notin \mathbf{P}(+1,-1) \), as follows from Corollary 4.

Remark 7 The class of 1-safe primes is not be confused with that of Sophie Germain primes ((Yates, 1991), (Ribenboim, 1991)), though they are related. Actually, if \( p = 2q + 1 \), with \( p \) and \( q \) prime integers, then \( p \) is a 1-safe prime and \( q \) is a Sophie Germain prime.

2.2. Mixed signatures

We introduce now the concept of mixed signature.

Definition 8 A signature \( (\varepsilon_1,...,\varepsilon_k) \) is said to be mixed whenever indices \( i, j \) exist such that, \( 1 \leq i < j \leq k \) and \( (\varepsilon_i, \varepsilon_j) = -1 \).

Proposition 9 The sets of mixed signature \( \mathbf{P}(\varepsilon_1,...,\varepsilon_k) \) are classified as follows:

1. If \( k = 2 \), then \( \mathbf{P}(+1,-1) = \{11\}, \mathbf{P}(-1,+1) = \{13\} \).
2. If \( k = 3 \), then \( \mathbf{P}(+1,+,1,-1) = \{23\} \) and all the other ones are empty.
3. If \( k = 4 \), then \( \mathbf{P}(+1,+1,+,1,-1) = \{47\} \) and all the other ones are empty.
4. If \( k > 5 \), then all the mixed signatures are empty.

The interesting observation here is that the set of mixed signatures is essentially empty.
2.3. Chains of safe primes

Definition 10 A chain (see, for example, (Forbes, 1999), (Teske and Williams, 2000)) of safe primes of length \( k \) is a sequence of prime integers \( p, p_i, \ldots, p_{k-1} \) such that, \( p = 2p_1 + \epsilon_1, p_i = 2p_{i-1} + \epsilon_i, 1 \leq i \leq k - 2 \), and \( \epsilon_i \in \{+1,-1\}, 1 \leq i \leq k - 1 \).

Remark 11 A prime \( p \) takes the first place in a \( k \)-length chain if and only if \( p \) belongs to \( P_{k-1}^+ \cup P_{k-1}^- \).

There exist a number of different kinds of prime chains: Cunningham chains, Shanks chains, etc. Only Cunningham chains are related to our notion of safe prime chains. A Cunningham chain (see, for example, (Teske and Williams, 2000), (Guy, 1994)) is a sequence of \( k \geq 2 \) prime integers \( p_i, \ldots, p_k \) such that \( p_{i+1} = 2p_i + \epsilon \), \( i = 1, \ldots, k - 1 \), \( \epsilon \in \{+1,-1\} \). Observe that if \( k = 2 \), then a Cunningham chain of length 2 is simply a pair \((q, 2q + 1)\), where \( q \) is a Sophie Germain prime.

Observe that, by definition, the last prime \( p \) of a Cunningham chain of length \( k \) and signature \( +1 \) (resp. \( -1 \)) verifies that \( p \in P_{k-1}^+ \) (resp. \( p \in P_{k-1}^- \)).

It is not known whether safe prime chains exist with arbitrary length \( k \), or equivalently if the sets \( P_k^+ \), \( P_k^- \) are not empty for every \( k \in \mathbb{N} \). S. Yates has proved in (Yates, 1991) that the maximal length for a Cunningham chain whose first term is \( p \) and \( \epsilon = 1 \) is \( p - 1 \).

Recent records for Cunningham chains have been obtained by P. Carmody and P. Jobling, for \( \epsilon = 1 \), and by T. Forbes (see (Forbes, 1999)) for \( \epsilon = -1 \). In both cases, the length of the longest chain is \( k = 16 \), and the first terms are \( p_1 = 810433818265726529159 \) for \( \epsilon = 1 \) and \( p_1 = 3203000719597029781 \) for \( \epsilon = -1 \). Both results are reported in (Caldwell, 2005).

3. CHAINS OF GENERALISED SAFE PRIMES

The case of positive signatures has received a detailed attention in the literature; see (Maurer, 1990), (Maurer, 1992), (Maurer, 1995), (Mihailescu, 1994), whereas the negative and mixed signatures have not been studied as far as we know, and, consequently, we shall deal with them below.

Following (Maurer, 1992) basically, we first relax the requirement in Definition 10 of chains of safe primes by considering chains of the form

\[
p = 2a_1 p_1 + \epsilon_1, \quad p_1 = 2a_2 p_2 + \epsilon_2, \ldots, \quad p_{k-1} = 2a_k p_k + \epsilon_k,
\]

for some positive integers \( a = (a_1, \ldots, a_k) \), which are assumed to be small enough with respect to the sizes of the prime integers \( p_1, \ldots, p_k \).
Let us first state a result that we will use later on. According to (Bateman and Stemmler, 1962), suppose that \( f_1, \ldots, f_s \) are distinct irreducible polynomials with integral coefficients and positive leading coefficients, and suppose \( F \) is their product. Let \( Q_{F(N)} \) be the number of positive integers \( j \) between 1 and \( N \) inclusive such that \( f_1(j), \ldots, f_s(j) \) are all primes.

Then for large \( N \) we have

\[
Q_{F(N)} \leq 2^s s! C(F) N (\log N)^{-s} + o(N (\log N)^{-s}).
\]

Moreover the constant \( C(F) \) is given by

\[
C(F) = \prod_p \left( 1 - \frac{1}{p} \right)^{-v(p)} \left( 1 - \frac{\omega(p)}{p} \right),
\]

and \( \omega(p) \) denotes the number of solutions of the congruence \( F(X) \equiv 0 \mod p \). The proof can be found in the reference above. As a result, the reference (Bateman and Horn, 1962) states the following

**Conjecture 12** With the same assumptions and notation as above, the following formula is conjectured:

\[
Q_{F(N)} \equiv h_{-1} h_{-2} \cdots h_{-s} C(F) \left( \frac{\log N}{2} \right)^{-s} du
\]

where \( h_1, \ldots, h_s \) stand for the degrees of the polynomials.

**Definition 13** The counting function \( \pi_{\bar{a}}^\bar{\epsilon} \) for \( k \)-safe prime numbers of signature \( \bar{\epsilon} = (\epsilon_1, \ldots, \epsilon_s) \) and vector \( \bar{a} = (a_1, \ldots, a_s) \) is defined to be the number \( \pi_{\bar{a}}^\bar{\epsilon}(x) \) of \( p \in \mathbb{P} \) such that \( p \leq x \), and \( p = 2a_1 p_1 + \epsilon_1, \ p_1 = 2a_2 p_2 + \epsilon_2, \ldots, \ p_{k-1} = 2a_k p_k + \epsilon_k \).

We first deal with the negative signatures.

### 3.1. Negative signature

Consider first the case when vector \( \bar{a} \) has equal components. Let us state the following result:

**Theorem 14** Under the **Conjecture 12**, if \( p_a \) is the least prime divisor of \( 2a - 1 \), then the density of the chains in the equation (1) for the vector with equal components \( a = \ldots = a_k = a \), and for \( \epsilon_1 = \ldots = \epsilon_k = -1 \), vanishes for all \( k \geq p_a - 1 \). In particular, this holds if \( 2a - 1 \) is a prime.

**Theorem 15** Let \( a = \ldots = a_k = a \), \( \epsilon_1 = \ldots = \epsilon_k = -1 \), and \( p_a \) be as in the previous Theorem, and let \( D(a) = \{ p \in \mathbb{P} : p | a \} \) be the set of prime divisors of \( a \in \mathbb{N} \). Assume \( k \leq p_a - 2 \). Then,

\[
\pi_{\bar{a}}^\bar{\epsilon}(x) = \frac{C(F)}{2a^k} \int_{\epsilon_{-1}}^{\epsilon_{-s}} \left( \ln \left( (2a)^s \left[ t + \frac{(2a)^s - 1}{2a - 1} \right] \right) \right)^{k-1} dt,
\]
where $l_{a,k} = \frac{2(2a)^{k+1} - 3(2a)^k + 1}{2a-1}$, and

$$C(F) = \prod_{p \in D(2a-1)} \left(1 - \frac{k+1}{p}\right) \cdot \prod_{p \in [2, \ldots, D(2a-1)]} \left(1 - \frac{1}{p}\right)^{-k} \cdot \prod_{p \in D(2a-1)} \left(1 - \frac{1}{p}\right)^{-k} \cdot \left(\frac{1 - \min(k+1, e(2a, p))}{p}\right).$$

$e(2a, p)$ being the order of $2a$ in $Z_p^\ast$.

We consider now the case when vector $\vec{a}$ has unequal components. In the case of a vector $\vec{a} = (a_1, \ldots, a_k)$ with unequal components the results are not so concrete as those in the previous section. First of all, we introduce some notations:

\[\alpha_{h,j} = 2^{h-j} \prod_{i=h-j+1}^{h} a_i, \quad 1 \leq h \leq k, 0 \leq j \leq h,\]  
\[\beta_h = \sum_{i=1}^{h} \alpha_{i,1} + 1, \quad 1 \leq h \leq k,\]  
\[N_p^k = \# \left\{0\right\} \cup \left\{\alpha_{h_0}\beta_h \mod p : \alpha_{h_0} \in Z_p^\ast, 1 \leq h \leq k\right\}.\]  

**Theorem 16** Let $p$ be a prime number. We consider the notations in equations (2)–(4). If either an index $1 \leq h \leq k$ exists such that, $\alpha_{h_0} \equiv 0 \mod p$ and $\beta_h \equiv 0 \mod p$, or $N_p^k = p$, then $\omega(p) = p$, and hence $C(F) = 0$.

**Remark 17** Let $p_a = \max \cup_{\alpha_0 \in \mathbb{Z}} D(\alpha_{h_0})$. If $p \geq p_a + 1$, then $\alpha_{h_0} \neq 0 \mod p$ for every $1 \leq h \leq k$.

**Theorem 18** With the same assumptions and notations as in **Theorem 16** and **Remark 17**, if

\[\{p \in P : 2 \leq p \leq k+1\} \subseteq \bigcup_{\alpha_0 \in \mathbb{Z}} (D(\alpha_{h_0}) - D(\beta_h))\]

then $\omega(p) < p$ for every $p \in \mathbb{P}$ and hence, $C(F) > 0$.

**Example 19** For $k = 6$ if we consider all the subsets

\[\{a_1, \ldots, a_6\} \subseteq \{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20\}\]

then the exact number of vectors $\vec{a} = (a_1, \ldots, a_6)$ fulfilling the hypotheses of **Theorem 18** is 304.

From the possible 6-tuples obtained above, we supply, as an example, the case when $\vec{a} = (5, 6, 14, 15, 18, 19)$. We, then, have

\[D(\alpha_{10}) = \{2, 19\}, D(\alpha_{20}) = \{2, 3, 19\}, \quad D(\alpha_{30}) = \{2, 3, 5, 19\},\]

\[D(\alpha_{40}) = D(\alpha_{50}) = D(\alpha_{60}) = \{2, 3, 5, 7, 19\}.\]

Similarly,

\[D(\beta_1) = \emptyset, D(\beta_2) = \{37\}, D(\beta_3) = \{11, 101\}, D(\beta_4) = \{13, 2393\},\]

\[D(\beta_5) = \{241, 1549\}, D(\beta_6) = \{41, 83, 1097\}.\]
Hence \( \{p \in P : 2 \leq p \leq 7 \} \subseteq \bigcup_{15 \leq k \leq 6} (D(\alpha_{\sigma_0}) - D(\beta_\sigma)) = \{2, 3, 5, 7, 19\} \).

For such values of \( \alpha \sigma \) the following ten numbers are safe primes of length \( k = 6 \) and negative signature:

\[
\begin{align*}
157213815390109, & \quad 448585785745309, \\
1025803746270109, & \quad 3371555445649309, \\
3933752569182109, & \quad 3948661711710109, \\
6787362449041309, & \quad 7090927318033309, \\
7228241182609309, & \quad 8204853725406109.
\end{align*}
\]

4. CONCLUSION

In this paper, we discuss the definition and basic properties of higher-order safe primes. We revisit the concept of chains of safe primes and report on the most recent records. We then focus our attention on negative and mixed signatures following the more relaxed definition given in (Maurer, 1992) for safe primes. This definition can be seen as a certain generalisation of the “classical” prime chains. For this type of primes, and based on the conjectures of Bateman-Horn (Bateman and Horn, 1962)—which in turn depend on those of Hardy-Littlewood (Hardy and Littlewood, 1922)—we provide explicit conditions to find primes fulfilling a particular set of conditions, thus allowing the design of algorithms to find them. Finally, we give some examples, obtained using these algorithms.

REFERENCES


