On the Security of Trapdoor Subgroups of $\mathbb{Z}_N^*$

Marc Joye
marc.joye@gmail.com

In a recent work, Park, Lee and Lee [1] present an identity-based encryption scheme and a companion signature scheme. Interestingly, their schemes work in RSA subgroups. The authors introduce two security assumptions to prove the security of their schemes. We show in this short note that, unfortunately, the assumptions do not hold.

ASSUMPTIONS. Consider PPT algorithm $\text{TRSAgen}$ which, on input a security parameter $1^κ$, generates an RSA modulus $N = pq$ where $p = 2p_1 + 1$ and $q = 2q_1 + 1$, with $p$, $q$, $p_1$, $q_1$ prime, an element $g \in \mathbb{Z}_N^*$ of order $p_1q_1$, and a length $ℓ$ for some $ℓ < |p_1q_1|$. We write $(N, g, l) \leftarrow \text{TRSAgen}(1^κ)$.

Assumption 1 (TSDH Assumption). With the previous notations, the $\tilde{q}$-TSDH assumption asserts that

$$\Pr \left[ A \left( N, g, g^x, g^{(x+r_0)y}, r_0, \{ (1/(x + r_i), r_i) \}_{i=1}^{\tilde{q}} \right) = g^y \right]$$

is negligible for any PPT algorithm $A$; the probabilities are taken over the experiment of running $(N, g, ℓ) \leftarrow \text{TRSAgen}(1^κ)$ and choosing at random $x, y \in \mathbb{Z}_{p_1q_1}$ and $r_0, r_1, \ldots, r_{\tilde{q}} \in \{0, 1\}^ℓ$.

Assumption 2 (TSEI Assumption). With the previous notations, the $\tilde{q}$-TSEI assumption asserts that

$$\Pr \left[ A \left( N, g, g^x, \{ (1/(x + r_i), r_i) \}_{i=1}^{\tilde{q}} \right) = 1/(x + r^*) \mid r^* \notin \{r_1, \ldots, r_{\tilde{q}}\} \right]$$

is negligible for any PPT algorithm $A$; the probabilities are taken over the experiment of running $(N, g, ℓ) \leftarrow \text{TRSAgen}(1^κ)$ and choosing at random $x \in \mathbb{Z}_{p_1q_1}$ and $r_1, \ldots, r_{\tilde{q}} \in \{0, 1\}^ℓ$.

ANALYSIS. Define the set $\mathcal{S} = \{ (\sigma_i, r_i) \}_{i=1}^{\tilde{q}}$ where $\sigma_i = 1/(x + r_i) \pmod{p_1q_1}$. From a pair of elements $(\sigma_1, r_1), (\sigma_2, r_2) \in \mathcal{S}$, it is easily checked that

$$\sigma_1\sigma_2(r_1 - r_2) + \sigma_1 - \sigma_2 \equiv 0 \pmod{p_1q_1}.$$ 

We can write $A := \sigma_1\sigma_2(r_1 - r_2) + \sigma_1 - \sigma_2$ as $A = 2^ℓA_0$ where $2^ℓ \| A$ and $A_0$ is odd. It is worth observing that $A_0 \propto p_1q_1$. Next we choose a random element $h \in \mathbb{Z}_N^*$ with Jacobi symbol $-1$. It follows then that

$$\gcd(h^{A_0} + 1 \mod N, N)$$

yields the factorization of $N$. Indeed, since $\left( \frac{h}{N} \right) = -1$, we can assume without loss of generality that $\left( \frac{h}{p} \right) = 1$ and $\left( \frac{h}{q} \right) = -1$. Hence, letting $α = A_0/p_1$ and $β = A_0/q_1$ in $\mathbb{Z}$, we have $h^{A_0} \equiv \left( \frac{h}{p} \right)^α \equiv 1 \pmod{p}$ and $h^{A_0} \equiv \left( \frac{h}{q} \right)^β \equiv -1 \pmod{p}$ since $β$ is odd. In turn, this implies that $\gcd(h^{A_0} - 1 \mod N, N) = p$ and $\gcd(h^{A_0} + 1 \mod N, N) = q$. Hence, we get $p_1q_1 = (p - 1)(q - 1)/4$.

Once $p_1q_1$ is known, an attacker against the TSDH assumption can recover $x$ from $(\sigma_1, r_1)$ as $x = (1/\sigma_1) - r_1 \pmod{p_1q_1}$ and next $z := g^x \mod N$ as $z = (g^{(x+r_0)y})^γ \mod N$ where $γ := 1/(x + r_0) \pmod{p_1q_1}$.

Similarly, an attacker against the TSEI assumption can first recover $x$ from $(\sigma_1, r_1)$ and then compute a new pair $1/(x + r^*) \pmod{p_1q_1, r^*}$ for any chosen $r^* \in \{0, 1\}^ℓ$. 
Extension. The authors of [1] more generally consider the case of RSA moduli of the form \( N = pq \) where \( p = 2p_1 \cdots p_I + 1 \) and \( q = 2q_1 \cdots q_J + 1 \) with \( p_1, \ldots, p_I \) and \( q_1, \ldots, q_J \) prime, and \( p_1q_1 \)-order subgroups of \( \mathbb{Z}_N^* \). In this case, an attacker, from two pairs \((\sigma_1, r_1)\) and \((\sigma_2, r_2)\) where \( \sigma_i = 1/(x + r_i) \mod p_1q_1 \) with \( i \in \{1, 2\} \), can compute \( \Lambda \) as

\[
\Lambda := \sigma_1 \sigma_2 (r_1 - r_2) + \sigma_1 - \sigma_2 .
\]

Again the main observation is that the so-defined \( \Lambda \) is a multiple of \( p_1q_1 \).

Given two integers \( a \) and \( b \) whose prime factorizations are \( a = \prod_{i \in \mathcal{I}} p_i^{e_i} \) and \( b = \prod_{j \in \mathcal{J}} p_j^{f_i} \) with \( e_i, f_i > 0 \), we define the operator \( \text{cop}_b(a) = \prod_{i \in \mathcal{I}, i \notin \mathcal{J}} p_i^{e_i} \). This can be obtained by successive GCDs:

\[
\delta \leftarrow \text{gcd}(a, b)
\]

**until** (\( \delta = 1 \)) **do**

\[
a \leftarrow a/\delta; \ \delta \leftarrow \text{gcd}(a, b)
\]

**return** \( a \)

Against the TSDH assumption, the attacker computes

\[
x' := (1/\sigma_1) - r_1 \mod \text{cop}_{\sigma_1}(A), \quad \gamma := 1/(x' + r_0) \mod \text{cop}_{(x' + r_0)}(A)
\]

and then recovers \( z := g^y \mod N \) as

\[
z = (g^{(x + r_0)y})^\gamma \mod N .
\]

For the TSEI assumption, an attacker needs to produce a pair \((1/(x + r^*), r^*)\) where the first component is computed modulo \( p_1q_1 \). The knowledge of a multiple of \( p_1q_1 \) (i.e., \( \Lambda \)) is not enough. However, by considering more than two pairs \((\sigma_i, r_i)\), the attacker can easily obtain \( p_1q_1 \); namely

\[
\Lambda' := \text{gcd}(\sigma_1 \sigma_2 (r_1 - r_2) + \sigma_1 - \sigma_2, \sigma_1 \sigma_3 (r_1 - r_3) + \sigma_1 - \sigma_3)
\]

is likely to be equal to \( p_1q_1 \) or a small multiple thereof (that allows the recovery of \( p_1q_1 \) by pulling out its small factors).

References