symmetry

MDPI

Article

On the Number of Witnesses in the Miller–Rabin Primality Test

Shamil Talgatovich Ishmukhametov *, Bulat Gazinurovich Mubarakov and Ramilya Gakilevna Rubtsova

Institute of Computational Mathematics and Information Technology, Kazan Federal University, Kremlevskya St. 35, Kazan 420008, Russia; mubbulat@mail.ru (B.G.M.); Ramilya.Rubtsova@kpfu.ru (R.G.R.)

* Correspondence: Shamil.Ishmukhametov@kpfu.ru

Received: 24 February 2020; Accepted: 28 March 2020; Published: 1 June 2020

Abstract: In this paper, we investigate the popular Miller–Rabin primality test and study its effectiveness. The ability of the test to determine prime integers is based on the difference of the number of primality witnesses for composite and prime integers. Let $W(n)$ denote the set of all primality witnesses for odd n. By Rabin's theorem, if n is prime, then each positive integer $a < n$ is a primality witness for n. For composite n, the power of $W(n)$ is less than or equal to $\varphi(n)/4$ where $\varphi(n)$ is Euler's Totient function. We derive new exact formulas for the power of $W(n)$ depending on the number of factors of tested integers. In addition, we study the average probability of errors in the Miller–Rabin test and show that it decreases when the length of tested integers increases. This allows us to reduce estimations for the probability of the Miller–Rabin test errors and increase its efficiency.

Keywords: prime numbers; primality test; Miller–Rabin primality test; strong pseudoprimes; primality witnesses

1. Introduction

The MillerRabin primality test is an algorithm that checks whether a given number is prime or composite. Its original version, due to Gary L. Miller, was deterministic and relied on the unproved extended Riemann Hypothesis [1]. Michael O. Rabin modified it to obtain a probabilistic algorithm [2].

Definition 1. *Let m be a positive integer represented as $m = 2^s \cdot u$ where u is odd. We introduce two auxiliary functions $bin(m) = s$ and $odd(m) = u$.*

Definition 2. *Let n be an odd natural, $n > 9$. An integer $a, 1 \le a < n$, is called a primality witness for n if it is co-prime to n and one of the following conditions holds:*

$$\begin{array}{l} 1.\ a^{odd(n-1)} \equiv 1 \bmod n, \\ 2.\ a^{odd(n-1)2^i} \equiv -1 \bmod n \text{ for some } i,\ 0 \le i < bin(n-1), \end{array} \tag{1}$$

(We replaced original Rabin's definition of the compositeness witnesses by the opposite relation). For generality, we count 1 and $n-1$ as primality witnesses and call them trivial witnesses since they satisfy (1) for any n.

Let $W(n)$ denote the set of all primality witnesses for n. The Rabin theorem [2] asserts that if number n is prime then each non-zero integer, $a < n$ is a primality witness for n, and therefore, the number of all witnesses $|W(n)| = n - 1$. For composite n, it satisfies inequality $|W(n)| \le \varphi(n)/4$ where $\varphi(n)$ is Euler's totient function. Since Rabin did not consider 1 as a witness, then he stated the strict inequality $|W(n)| < \varphi(n)/4$.

Later, Gary Miller [1] developed a primality test that takes any integer a, $1 < a < n$, checks if a is not a factor of n (otherwise, n is trivially composite), and whether a is a primality witness for n, that is, lies in the set $W(n)$. If the answer is positive, then n is probable prime with probability exceeding $3/4$. If we need in a more exact result, we should repeat this procedure several times taking different numbers $a < n$.

The researchers refer to this algorithm as to the Miller and Rabin primality test. We abbreviate it to *MR test*.

Definition 3. *Parameters a which are used in Miller's algorithm are called* bases. *They are chosen randomly from interval* $[1; n-1]$. *If, for a given odd integer, n relation (1) holds at a base a, we say, n passes the MR test at base a. Otherwise, we call a a compositeness witness for n and deduce that n is certainly composite.*

The probability of error after k successful iterations becomes less than $1/4^k$. The only type of error in the Rabin' procedure is defining a composite integer as prime.

More details on the Miller–Rabin test can be found in Chapter 3 of text-book [3] by Crandall and Pomerance. We abbreviate Miller–Rabin test as MR test.

Definition 4. *Composite integers qualifying by MR test as probable prime at a base a are called strong pseudoprimes relative to base a. Composite integers being probably prime relative to all a from a set A of bases are called strong probable prime relative to set of bases A.*

Investigation of pseudoprime integers has a long history in the Computational Number Theory. We outline main advantages in this direction in the next section.

2. Some History Remarks

Fist attempts to find fast primality algorithms were based on Fermat's Little Theorem asserting that for prime n and for any positive integer a, the following relation holds

$$a^n \equiv a \bmod n \tag{2}$$

Indeed, many composite integers do not satisfy (2) and can be discarded after the first check. Composite n that satisfy (2) are called Fermat pseudoprimes relative to base a.

It is important to note that all strong pseudoprimes relative to a base a are also Fermat pseudoprimes relative to a.

We can decrease the number of false decisions by Fermat's test by checking the relation (2) with several different a. However, this does not allow us to completely avoid false conclusions since so-called Carmichael numbers exist.

Integer n is called a Carmichael number if it satisfies (2) for all a. Carmichael numbers appear relatively rarely and the least Carmichael number is $561 = 3 \cdot 7 \cdot 11$. It is known that Carmichael numbers are exactly those integers which satisfy Korselt's criterion:

Korselt Criterion (1899). A positive compositeinteger n is a Carmichael number if and only if n is square-free, and for all prime divisors p of n, it is true that $p-1|n-1$.

One of the interesting problems is to find for a given odd integer n the least witness. In 1994 Alford, Granville and Pomerance proved [4] that such witnesses exceed $(\log n)^{1/(3\log\log\log n)}$ for infinitely many n. We also show that there are finite sets of odd composites which do not have a reliable witness, namely a common witness for all of the numbers in the set.

MR test discards a Carmichael number n, if the base was chosen from $[1; n-1] \backslash W(n)$.

Let us fix a base a and let n_a be a least composite integer that the MR Test accepts at the base a. Then, any odd $n < n_a$ for which a is a primality witness, is definitely prime. This means that when we know n_a, we can definitely check any $n < n_a$ for primality using only one round of the MR procedure. The corresponding integer n_a is small. But if we take a set A of several different bases a and find a

least composite n_A for which all $a \in A$ are primality witness, this n_A can be very large. Candidates for bases a can be any positive integers that are not squares. However, historically, candidates for special bases are chosen from the set of primes.

Let P_k denote the set of the first k primes $P_k = \{2, 3, 5, 7, \ldots, p_k\}$, and let ψ_k be a least strong pseudoprime relative to P_k for a $k \geq 1$. Function ψ_k is well defined and is exponentially computable. Its computation began already 40 years ago.

First four values of ψ_k have been found by C. Pomerance, J. Selfridge, and S.Waggstaff [5] in 1980.

A systematic calculation of ψ_k for larger k has been initiated by J. Jaeschke [6] who elaborated basic algorithms helpful for searching for strong pseudoprimes of different forms. In 1993 Jaeschke calculated ψ_k for $5 \leq k \leq 8$ and proposed upper bounds for ψ_k at $9 \leq k \leq 11$.

F. Arnault in papers [7,8] described another algorithm to search for Carmichael numbers and strong pseudoprimes integers.

Jaeschke' hypothesis have been improved in 2001 by Z. Zang [9] who constructed a lesser 19-digits decimal integer $Q_{11} = 3825123056546413051$ bounding above ψ_{11}. Z.Zang conjectures that values ψ_k for $9 \leq k \leq 11$ are equal to each other and coincide with Q_{11}.

In 2012 J. Jiang and Y. Deng [10] confirmed Zang's Hypothesis by showing that $Q_{11} = \psi_9 = \psi_{10} = \psi_{11}$.

The last record is reached by J. Sorenson and J. Webster [11] in 2016 . They found ψ_{12} and ψ_{13}, where $\psi_{13} = 3317044064679887385961981 \approx 3.3 \cdot 10^{24}$. So at the moment we can successfully determine prime integers less than $3.3 \cdot 10^{24}$ by only 13 rounds of the MR test. But this bound is much less than integers used in Cryptography. For example, DSS algorithm uses prime integers of length 256 bits (≈80 decimal digits).

Another branch of investigations in connected with the problem of distribution of Fermat pseudoprimes and strong pseudoprimes. Let $F(n)$ denote set

$$F(n) = \{a \bmod n : \ a^{n-1} \equiv 1 \bmod n\}.$$

Clearly, $F(n) \supseteq W(n)$.

In 1985 P. Erdos and C. Pomerance [12] studied an asymptotic behavior of average function

$$A(x) = \frac{1}{x} \sum_{n \leq x}{}' |F(n)|$$

where sum is counted over odd integers. They showed using complex number-theoretical calculations that $A(x)$ is a growing function bounded below by $x^{15/23}$.

Our average function $Avg(x)$ looks close to $A(x)$ but we show that for almost all composite n $W(n)$ consists of only two elements 1 and $n-1$ and function $Avg(x)$ tends to zero with x tending to infinity.

Average number of errors in the MR test was also studied in 1993 by I. Damgard, P. Landrock and C Pomerance. In paper [13] they studied an average probability of the false decision by the MR test in the following procedure:

Fix $k > 0$ and $t > 0$ and choose randomly k-bit odd integer n. Check it with t rounds of MR test with randomly chosen bases from $[1; n-1]$. If n was discarded during the procedure (that is, found $a \notin W(n)$), take another n. Continue until n was found passed t rounds. Let $p_{k,t}$ be the probability that the procedure returns a composite integer.

The authors found explicit upper bounds for various k and t. In particular they proved that $p_{k,1} \leq k^2 4^{2-\sqrt{k}}$ for $k \geq 2$. Their results show that the probability of false decisions of the MR test depends on the length of tested numbers and it decreases if the length of the numbers increases.

3. Counting Number of Witnesses

In this section we deduce exact formulas for the number of primality witnesses for different types of composite integers.

We begin our investigation with a little proposition improving Rabin's estimate.

Theorem 1. *If $a \in W(n)$, then $n - a \in W(n)$.*

Proof. Let $k = ord_n(a)$. If k is odd, then $a^{odd(n-1)} \bmod n = 1$, and $(n-a)^{odd(n-1)} \equiv -1 \bmod n$, therefore, $n - a$ is also a witness.

If k is even, then $a^{k/2} \equiv -1 \bmod n$. If $k/2$ is even, then $(n-a)^{k/2} \equiv a^{k/2} \equiv -1 \bmod n$, and $(n-a)$ is a witness.

Finally, if $k/2$ is odd, then $(n-a)^{k/2} \equiv -a^{k/2} \equiv 1 \bmod n$. Since $k/2 \mid odd(n-1)$, then $a^{odd(n-1)} \equiv 1 \bmod n$, and $(n-a)$ again is a witness.

This completes the proof. □

Corollary 1. *(The Improved Rabin Theorem). Let n be a natural, and A be an arbitrary set of bases less than n, co-prime to n, such that for any $a \in A$, $n - a$ is not in A. If all bases $a \in A$ are primality witnesses of n, then n is probable prime with probability of error less than or equal to $1/16^k$.*

Indeed, when we found a primality witness a for integer n, we get two primality witnesses for n, namely, a and $n - a$. So, this reduces the probability of error by a factor of $4^2 = 16$.

Let $N_w(n) = |W(n)|$ be the power of number of primality witnesses $W(n)$. As mentioned earlier, for prime n $N_w(n) = n - 1$, and for composite n $N_w \leq \varphi(n)/4$.

Below we estimate function $N_w(n)$ more exactly. First we formulate a theorem restricting possible witnesses for a composite n.

Theorem 2. *Let $n = u \cdot v$ for co-prime factors u and v (possibly, composite), and $a \in W(n)$. Then,*

$$\begin{array}{l} 1.\ ord_u(a) \mid GCD(\varphi(u),\ (u - \varphi(u))v - 1), \\ 2.\ ord_v(a) \mid GCD(\varphi(v),\ (v - \varphi(v))u - 1), \\ 3.\ bin(ord_u(a)) = bin(ord_v(b)). \end{array} \tag{3}$$

Proof. 1. Since a is a primality witness for n then $a^{n-1} \equiv 1 \bmod n$ and $a^{n-1} \equiv 1 \bmod u$. Besides, $n - 1 = uv - 1 = \varphi(u)v + (u - \varphi(u))v - 1$, so

$$1 \equiv a^{n-1} \equiv a^{\varphi(u)v + (u-\varphi(u))v - 1} \equiv a^{(u-\varphi(u))v-1} \bmod u,$$

since $a^{\varphi(u)} \equiv 1 \bmod u$ by Euler's Theorem.

2. By symmetry.

3. If $ord_u(a)$ is odd, then $a^{odd(n-1)} \equiv 1 \bmod n$ (otherwise, a satisfies the second clause of the MRT, and $ord_u(a)$ should be even). Then $a^{odd(n-1)} \equiv 1 \bmod v$ and $ord_v(a)$ is odd.

If $bin(ord_u(a)) = i$ for $0 < i < bin(n-1)$, then a is a witness by second clause of the MRT, so $a^{odd(n-1)2^{i-1}} \equiv -1 \bmod n$, $a^{odd(n-1)2^{i-1}} \equiv -1 \bmod v$, and $a^{odd(n-1)2^i} \equiv 1 \bmod v$, so $ord_v(a) = odd(n-1)2^i$ and $bin(ord_v(a))$ is equal to i.

The theorem is proved. □

Example 1. *Let $n = 15 \cdot 19 = 285$, and $a \in W(n)$. By Theorem 2:*

$$\begin{array}{l} 1.\ ord_u(a) \mid GCD(\varphi(u),\ (u - \varphi(u))v - 1) = GCD(8, 132) = 4, \\ 2.\ ord_v(a) \mid GCD(\varphi(v),\ (v - \varphi(v))u - 1) = GCD(18, 14) = 2, \\ 3.\ bin(ord_u(a)) = bin(ord_v(b)). \end{array}$$

So, possible a satisfies $(ord_u(a), ord_v(a)) = (1,1)$, *or,* $(ord_u(a), ord_v(a)) = (2,2)$, *so* $n = 285$ *has only trivial witnesses* 1 *and* $n-1$.

Theorem 3. *Let* $n = p^k$ *be a degree of prime* p, *then* $N_w(n) = p-1$.

Proof. Let a be a witness for $n = p^k$, then $ord_a(n) \mid GCD(\varphi(n),\ n-1) = GCD(p^{k-1}(p-1),\ p^k-1) = p-1$.

Besides, any a satisfying $a^{p-1} \bmod n = 1$ is a witness of n. Indeed, let $a^{p-1} \bmod n = 1$. Then, $m = ord_n(a)$ is a factor of $n-1 = p^k-1$. Let $n-1 = 2^s \cdot t$ for odd t, therefore, $m = 2^{s_1} \cdot t_1$, where $s_1 \le s$ and t_1 is a factor of t.

If $s_1 = 0$, then $a^{t_1} \bmod n = 1$, $a^t \bmod n = 1$ and a is a witness by the first clause of the MRT. Otherwise, let $0 \le r \le s_1$ be such that $a^{t_1 2^r} \equiv -1 \bmod n$. Then $a^{t2^r} \equiv -1 \bmod n$ and a is a witness by the second clause of the MRT. This completes the proof. □

We call integer n *semiprime* if it is a product of two distinct primes $n = pq,\ p < q$. Semiprimes are close to primes, and we prove below that they have a maximal number of primality witnesses among composite numbers.

Theorem 4. *Number of witnesses of semiprime* $n = pq$ *is equal to*

$$N_w(pq) = (odd(d))^2 \cdot (4^{bin(d)} + 2)/3, \tag{4}$$

where $d = GCD(p-1,\ q-1)$.

We begin with example of application of this formula.

Example 2. *Let* $n = 11 \cdot 31 = 341$. *Then* $d = GCD(p-1,\ q-1) = 10 = 5 \cdot 2^1$, $odd(d) = 5$, $s = bin(d) = 1$. *By the theorem,*

$$N_w(31) = 5^2 \cdot (4+2)/3 = 50.$$

Proof. Let $d = GCD(p-1,\ q-1)$. Applying Theorem 2 to $n = pq$ we obtain

1. $ord_p(a) | d,\ \ ord_q(a) | d$,
2. $bin(ord_u(a)) = bin(ord_v(b))$.

We distribute all n-witnesses a into $s+1$ classes W_i, $0 \le i \le s$, where class W_i consists of a with $bin(ord_p(a)) = bin(ord_q(a)) = i$.

Class W_0 contains such a that both $ord_p(a)$ and $ord_q(a)$ are odd. Let $a \in W_0$, and $(i,j) = (ord_p(a), ord_q(a))$. Numbers i and j are factors of $u = odd(d)$ by the choice of a. Conversely, each integer $a < n$ satisfying $ord_p(a) \mid u,\ \ ord_q(a) \mid u$, is a witness of n and lies in W_0.

Let fix a pair $(i,j),\ i|d, j|d$. By Euler's theorem, in Z_p there are exactly $\varphi(i)$ elements of multiplicative order i, and in Z_q there are $\varphi(j)$ elements of multiplicative order j, so, there exist exactly $\varphi(i) \cdot \varphi(j)$ pairs $(x,y), 0 < x < p,\ 0 < y < q$, such that $(ord_p(x), ord_q(y)) = (i,j)$. But for each such pair (x,y) there exists a unique $a < n$ with $(a \bmod p, a \bmod q) = (x,y)$, so there is a injective correspondence between witnesses a of n with odd orders $ord_p(a)$, $ord_q(a)$, and pairs (x,y) with $x|u$, $y|u$. Therefore, the power of W_0 is equal to

$$|W_0| = \sum_{x|u,\ y|u} \varphi(x) \cdot \varphi(y) = \left(\sum_{x|u} \varphi(x)\right)\left(\sum_{y|u} \varphi(y)\right) = u^2,$$

since by a known theorem of Euler for any natural m $\sum_{v|m} \varphi(v) = m$.

The next class W_1 has the same power u^2 since is consists of witnesses a with $bin(ord_p(a)) = bin(ord_q(a)) = 1$, and

$$|W_1| = \sum_{x|d,\ y|d} \varphi(2x) \cdot \varphi(2y) = u^2,$$

since $\varphi(2z) = \varphi(z)$ for odd z.

The power of class W_i is equal to

$$\sum_{x|d,\ y|d} \varphi(2^i x) \cdot \varphi(2^i y) = 4^{i-1}u^2.$$

Therefore, the number of all witnesses $N_w(n) = u^2(1+1+4+\ldots+4^{s-1}) = u^2 \cdot (4^s+2)/3$. This completes the proof. □

Corollary 2. *(Rabin's theorem for semiprimes). The number of witnesses of $n = pq,\ p \le q$, is less or equal to $\varphi(n)/4$.*

Proof. If $p = q$, then $N_w(n) = p - 1$ by Theorem 3, and $\varphi(n)/4 = p(p-1)/4$, so $N_w(n) < \varphi(n)/4$ at $p \ge 5$.

Let $p < q$. Ratio $N_w(n)/n$ reaches its maximum when $GCD(p-1; q-1) = p-1$, $q = 2p-1$, and $bin(p-1) = 1$. Indeed, $odd(n)$ is diminishing in two times when $bin(p-1)$ is added by 1, and the whole expression in (4) becomes less. Then, $max\ odd(d) = (p-1)/2$, so

$$max\ N_w(pq) = N_w(p(2p-1)) = \frac{(p-1)^2}{2} = \frac{\varphi(n)}{4}.$$

□

Example 3. *Let $n = 7 \cdot 13 = 91$. $N_w(91) = 3^2 \cdot 2 = 18 = \varphi(91)/4$.*

Now we study function $N_w(n)$ at products of k distinct primes. The general result for such products is formulated below:

Theorem 5. *Let $n = p_1 \cdot p_2 \cdot \ldots p_k$ be the product of k distinct primes. Then*

$$N_w(n) = u_1 \cdot u_2 \cdot \ldots \cdot u_k \cdot \left(1 + \frac{2^{ks}-1}{2^k-1}\right), \text{ where}$$

$$s = min\{bin(d_1), bin(d_2), \ldots, bin(d_k)\},\ \ d_i = GCD\left(p_i - 1; \prod_{j \ne i} p_j - 1\right),$$

$u_i = odd(d_i)$.

Let us begin with an example $n = 7 \cdot 13 \cdot 31 = 2821$. The corresponding restrictions are listed below:

1. $ord_p(a) \mid d_1 = GCD(p-1; qr-1) = 6,\ \ u_1 = 3,$
2. $ord_q(a) \mid d_2 = GCD(q-1; pr-1) = 12,\ u_2 = 3,$
3. $ord_r(a) \mid d_3 = GCD(r-1; pq-1) = 30,\ u_3 = 15,$
4. $bin(ord_p(a)) = bin(ord_q(b)) = bin(ord_r(b))$.

Since $s = min\{bin(d_1), bin(d_2), bin(d_3)\} = min\{1,\ 2,\ 1\} = 1$, we obtain

$$N_w(2821) = 3 \cdot 3 \cdot 15 \left(1 + \frac{2^3-1}{2^3-1}\right) = 270$$

(compare with $\varphi(n)/4 = 6 \cdot 12 \cdot 30/4 = 540$).

Proof. Let $u_i = odd(d_i)$ and k-tuple $(x_1, x_2, \ldots, x_k)$ contains components $x_i \mid u_i$, $1 \le i \le k$. There are $\varphi(x_1) \cdot \ldots \cdot \varphi(x_k)$ witnesses of n with $ord_{p_i}(a) = x_i$ for $1 \le i \le k$. So,

$$|W_0| = \sum_{(x_1,x_2,\ldots,x_k),x_i|u_i} \varphi(x_1) \cdot \ldots \cdot \varphi(x_k) =$$

$$= \left(\sum_{x|u_1} \varphi(x)\right) \cdot \left(\sum_{x|u_2} \varphi(x)\right) \ldots \left(\sum_{x|u_k} \varphi(x)\right) = u_1 \cdot u_2 \cdot \ldots \cdot u_k.$$

As in the previous theorem, the power of class W_1 is equal to power of $W_0 = u_1 \cdot u_2 \cdot \ldots \cdot u_k$, while the power of the each further class W_{i+1} is equal to the power of the previous one multiplied by $\varphi(2^k) = 2^{k-1}$ since each additive $\varphi(2^i x_1) \cdot \ldots \cdot \varphi(2^i x_k)$ in the previous class corresponds to additive $\varphi(2^{i+1} x_1) \cdot \ldots \cdot \varphi(2^{i+1} x_k)$ and their ratio r_i is

$$r_i = \frac{\varphi(2^{i+1} x_1) \cdot \ldots \cdot \varphi(2^{i+1} x_k)}{\varphi(2^i x_1) \cdot \ldots \cdot \varphi(2^i x_k)} = 2^k.$$

The proof is complete. □

4. Frequency Function

In this part we introduce a notion of *frequency function* that characterizes the probability to find at one attempt a primality witness for a given integer n.

Let define frequency function $Fr(n)$ as follows

$$Fr(n) = \frac{N_w(n)}{\varphi(n)}.$$

According to Rabin's theorem, $Fr(n) = 1$ for prime n, and $Fr(n) \le 1/4$ for composite n. We study distribution of values $Fr(n)$ for semiprime integers $n = pq,\ p < q$.

1. We begin our research with case $q - 1 = k(p - 1)$ for $k \ge 2$. Numbers of this type appear frequently among strong pseudoprimes. Let rewrite p and q in form $p = 2^s u + 1,\ q = 2^s ku + 1$, where u is odd, $s \ge 1$, and consider different s:

Case 1. $s = 1,\ u = odd(d) = (p-1)/2,\ \ N_w(pq) = 2u^2 = (p-1)^2/2,$

$$Fr(n) = \frac{(p-1)^2/2}{(p-1)(q-1)} = \frac{2u^2}{2u \cdot 2ku} = \frac{1}{2k}.$$

Function $Fr(n)$ reaches its maximum $1/4$ at $k = 2$: $(p, q) = (2u + 1,\ 4u + 1)$. Since, both p and q are prime then $u \equiv 0 \bmod 3$, so $(p, q) = (6t + 1, 12t + 1)$, $t \ge 1$. Such pairs form a sequence

$$(7,\ 13),\ (19,\ 37), (31,\ 61),\ (37,\ 73), \ldots.$$

Case 2. $s = 2,\ u = odd(d) = (p-1)/4,\ \ N_w(pq) = 6u^2$, *and*

$$Fr(n) = \frac{6u^2}{(p-1)(q-1)} = \frac{6u^2}{4u \cdot 4ku} = \frac{3}{8k}.$$

Maximum of $Fr(n)$ is now $3/16 = 0.1875$ at $k = 2$.

Case 3. $s \geq 1$, *At arbitrary s we have*

$$Fr(n) = \frac{(1+(4^s-1)/3)u^2}{(p-1)(q-1)} = \frac{(1+(4^s-1)/3)u^2}{2^s u \cdot 2^s k u} = \frac{1}{3ku^2 \cdot 2^{2s-1}} + \frac{1}{3k}.$$

Thus, function $Fr(n)$ *at semiprimes* $n = pq$, $q-1 = k(p-1)$, *is located in the interval*

$$\frac{1}{3k} < Fr(n) \leq \frac{1}{2k},\ k \geq 2. \tag{5}$$

2. Now, we turn to a common case $n = pq$:

$$p = 1 + k_1 u,\ q = 1 + k_2 u,\ GCD(k_1, k_2) = 1,\ u = t2^s,\ t\ odd.$$

For such n

$$N_w(n) = t^2(4^s+2)/3,\ \varphi(n) = k_1 k_2 t^2 4^s,\ Fr(n) = \frac{4^s+2}{3k_1 k_2 \cdot 4^s}.$$

So,

$$\frac{1}{3k_1 k_2} < Fr(n) \leq \frac{1}{2k_1 k_2}$$

Conclusion. Function $Fr(n)$ at semiprimes $n = pq$ depends mostly on values k_1 and k_2 in representation $p = k_1 u + 1$, $q = k_2 u + 1$. $Fr(n)$ takes maximal values close to $1/4$ only at small k_1 and k_2. This completely corresponds to experimental data. Among values ψ_k the most expected are pseudoprimes of form $u = (u+1)(2u+1)$ with minimal values $k_1 = 1$ and $k_2 = 2$.

An important question connecting with efficiency of MRT is the average frequency of witnesses for composite numbers. As earlier, we study this problem for semiprime integers.

Let fix any prime p and a board B. We count average frequency of integers pq, $q > p$, $pq \leq B$. For convenience, we assume that $B = p(p + (p-1)k)$ for a positive $k \in \mathbf{Z}$.

For simplicity we explain all deductions at example $p = 11$. Every prime q has $d = GCD(p-1,\ q-1)$ equal either 2, or 10.

Let $d = 10$. Corresponding q lie in the set $\{21, 31, 41, 51, 61, 71, 81, 91, 101, \ldots, 10k+11\}$, where $10k + 11 = B/p$. Each third integer in the sequence is a multiple of 3, some others are multiples of 7, 11 etc. Since q should be prime we need to remove them from the sequence. The rest consists of integers

$$Q_B = \{31, 41, 61, 71, 101, 113 \ldots\}. \tag{6}$$

We assume that primes $q \in Q_B$ are distributed uniformly in the interval $[1,\ B/p]$. Then the average frequency can be estimated as

$$Avg(Fr(n)) \approx \frac{1}{k}\left(\frac{1}{4} + \frac{1}{6} + \ldots + \frac{1}{2k}\right) = \frac{1}{2k}\left(1 + \frac{1}{2} + \frac{1}{3} + \ldots + \frac{1}{k}\right)$$

(we remind that $Fr(p(i(p-1)+p) = 1/2(i+1)$).

The expression in the last brackets is a partial sum of the Harmonic Series. Its value is

$$\sum_{i=1}^{k} \frac{1}{i} < \sum_{i=1}^{k+1} \frac{1}{i} = \ln k + \gamma + \varepsilon_n,$$

where $\gamma = 0.5772...$ is the Euler—Mascheroni constant and $\lim_{k\to\infty} \varepsilon_n = 0$. Constant γ and additive ε_n can be ignored so

$$Avg(Fr(n)) < \frac{\ln k}{2k}$$

Since $(p-1)k+1 = B/p$, then $k > B/p^2 - 1$ and $\ln k < \ln B$, so

$$Avg(Fr(n)) < \frac{\ln B}{2(B-p^2)} \cdot p^2 \tag{7}$$

Let us move now to primes q of type $d = GCD(p-1, q-1) = 2$. They lie in the sequence

$$q \in \{13,\ 15,\ 17,\ 19,\ 23,\ ,25,\ 27,\ 29, \ldots,\ 2m+1\}$$

where $2m+1 = B/p$, $q = 2i+1$, $GCD(i,5) = 1$. When we remove composite integers, the rest contains at least half members.

Integers $n = pq$ with $GCD(p-1, q-1) = 2$ have only trivial witnesses 1 and $n-1$ so their frequency function takes values

$$Fr(n) = \frac{2}{(p-1)(q-1)}.$$

Assuming that such n are distributed uniformly in the interval $[p^2; B]$ we estimate the average frequency by expression

$$Avg(Fr) \approx \left(2 \sum_{p \le k \le m} \frac{1}{(p-1)(2k+1)}\right) / \left(\frac{2m+1-p}{2}\right) <$$

$$\frac{4}{(2m+1-p)(p-1)} \cdot \frac{1}{2} \cdot \sum_{i=(p+1)/2}^{m} \frac{1}{i} < \frac{2}{(2m+1-p)(p-1)} \cdot \ln m$$

Substituting in the last expression $2m+1 = B/p$ we get

$$Avg(Fr) < \frac{2p \ln B}{(B-p^2)(p-1)} \tag{8}$$

Expressions (7) and (8) give upper bounds for two types of integers $n = pq$. In the second case the estimation is lesser so average estimation for the united class of all $n = pq \le B$, $p < q$, can be set by the upper bound of (7). This assertion does not depend on a special $p = 11$ so we can state the following theorem.

Theorem 6. *Let p be a prime and B satisfy $B > p^2$. Then the average frequency of witnesses in the class of semiprimes $n = pq \le B$, $q > p$, has an upper bound*

$$Avg(Fr(n)) < \frac{p^2 \ln B}{2(B-p^2)}$$

Note than limit of the average function is 0 as $B \to \infty$. This explains the phenomenon that the number of false conclusions in the Miller–Rabin test decreases when length of tested integers increases.

5. Numbers with Maximal Frequency of Witnesses

In this section we study composite n with maximal frequency $Fr(n) = 1/4$. Let $n = p_1 p_2 \ldots p_k$ be the product of k different primes.

We begin with case $k = 2$. As we see from the previous section, integers $n = pq$ have maximal frequency only in case when $q = 2p-1$. Such pairs appear comparatively often, and their quantity is diminishing together with their size.

Table 1 contains number of semiprimes with maximal frequency in intervals $[(i-1) \cdot 10^5;\ i \cdot 10^5;]$, $1 \le i < 10$.

Table 1. Distribution of semiprimes with maximal frequency below 10^6.

1	2	3	4	5	6	7	8	9	10
670	494	448	412	424	386	393	358	370	343

Case $k = 3$ is more interesting. In order function $Fr(pqr)$ reached its maximum = 0.25, we need satisfaction of four requirements:

$$\begin{array}{l} 1.\ GCD(p-1; qr-1) = p-1, \\ 2.\ GCD(q-1; pr-1) = q-1, \\ 3.\ GCD(r-1; pq-1) = r-1. \\ 4.\ bin(p-1) = bin(q-1) = bin(r-1) = 1. \end{array} \tag{9}$$

Such triples exist, and an example of it was already given in Rabin's paper [2] $n = 487 \cdot 1531 \cdot 2683 = 2000436751$. Rabin himself estimated $Fr(n)$ as 0.2493, but the difference is due to the fact that he did not include 1 in the list of witnesses.

Such triples appear much more seldom and have a form

$$n = (2k_1 + 1)u \cdot (2k_2 + 1)u \cdot (2k_3 + 1)u \text{ for } u \in N.$$

We arranged the search of such triples at a computer and found 160 such integers not exceeding $2 \cdot 10^{14}$. The least triple we found is

$$n = 19 \cdot 199 \cdot 271 = 1024651.$$

The largest found triple has a form $n = (u + 1)(3u + 1)(5u + 1)$ at $u = 24102$:

$$n = 24103 \cdot 72307 \cdot 120511 = 21002\,84533\,02331.$$

Let us study the form $\langle u, 3u, 5u\rangle$ and find restrictions on u in order to $n = (u + 1)(3u + 1)(5u + 1)$ satisfies first 3 conditions of (9). The first requirement is satisfied automatically. The second and third requirement are listed below:

$$(3u + 1) - 1 \mid (u + 1)(5u + 1) - 1 \ \rightarrow\ u \equiv 0 \bmod 3.$$

$$(5u + 1) - 1 \mid (u + 1)(3u + 1) - 1 \ \rightarrow\ 3u + 4 \equiv 0 \bmod 5,$$

so $u = 6 + 15t$ for $t \geq 1$. If we add requirements $p \equiv q \equiv r \equiv 3 \bmod 4$ we obtain

$$15t + 7 \equiv 3 \bmod 4 \ \rightarrow\ t \equiv 1 \bmod 4,\ u = 6 + 15(1 + 4t_1) = 21 + 60t_1.$$

Let now consider products of k primes where $k \geq 4$. The maximum of frequency of such products is $1/2^{k-1}$, since it is reached when for any $i \leq k$ $(p_i - 1)/2$ is odd, and $(p_i - 1) \mid \prod(p_{j\neq i} - 1)$. Then,

$$Fr(p) = 2 \cdot \prod_{i=1}^{k} \frac{p_i - 1}{2} = \frac{\varphi(n)}{2^{k-1}}.$$

A quick search of tuples $n = pqrt$ below 10^{12} gave 70 examples of them. The least 4-tuple was

$$n = 19 \cdot 31 \cdot 127 \cdot 547 = 40917241,$$

while the largest was

$$n = 19 \cdot 127 \cdot 14071 \cdot 29347 = 99\,64281\,70081.$$

Some computational results on distribution of strong semiprime integers can be found in [14].

6. Conclusions

In this section we will summarize the main results of the paper.

1. We found exact formulas for the number of witnesses for composite n with different number of factors.
2. We introduced the frequency function $Fr(n)$ characterizing the probability to find at one attempt a primality witness for a given n and found exact bounds for distribution of this function for semiprime integers n.
3. Like as Damgard, Landrock, and Pomerance in [13], we studied an average values of $Fr(n)$ at intervals $[1; x]$ for semiprime integers $n = pq,\ n \leq x$,with fixed p and showed that it bounded above by $p^2 \log x/2(x - p^2)$.
   Since such integers have maximal values of $F(n)$ among all composites, this opens a way in future investigations to find exact upper bounds for average values of frequency function among all k-bit odd integers for any k.
4. Finally, we described possible forms of composites with maximal values of frequency function for products of k distinct primes at $k \geq 2$ and using computer calculations found their examples and their quantity at initial intervals of set of all naturals.

Author Contributions: S.T.I. gave impetus to the research and proved Theorems 1 and 2. B.G.M. proved other theorems and propositions, and R.G.R. developed software for testing results. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by RFBR grant 18-47-16005. This investigation was supported by the grant of Scientific and Educational Mathematical Center of the Volga Federal District, agreement No. 075-02-2020-1478.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Miller Gary, L. Riemann's Hypothesis and Tests for Primality. *J. Comput. Syst. Sci.* **1976**, *13*, 300–317. [CrossRef]
2. Rabin, M. Probabilistic algorithm for testing primality. *J. Number Theory* **1980**, *12*, 128–138. [CrossRef]
3. Crandall, R.; Pomerance, C. *The Prime Numbers: A Computational Perspertive*, 2nd ed., Springer: Berlin, Germany, 2005; 604p.
4. Alford, W.R.; Granville, A.; Pomerance, C. On the difficulty of finding reliable witnesses. In *Algorithmic Number Theory*, First Internat. Symp., ANTS-I; Lecture Notes in Computer Science; Springer: Berlin/Heidelberg, Germany, 1994; p. 116.
5. Pomerance, C.; Selfridge, J.L.; Wagstaff, S.S., Jr. The pseudoprimes to $25 \cdot 10^9$. *Math. Comput.* **1980**, *35*, 1003–1026.
6. Jaeschke, G. On Strong Pseudoprimes to Several Bases. *Math. Comput.* **1993**, *61*, 915–926. [CrossRef]
7. Arnault, F. Rabin-Miller primality test: composite numbers which pass it. *J. Symb. Comput.* **1995**, *64*, 355–361. [CrossRef]
8. Arnault, F. Constructing Carmichael numbers which are strong pseudoprimes to several bases. *J. Symb. Comput.* **1995**, *20*, 151–161. [CrossRef]
9. Zhang, Z. Finding strong pseudoprimes to several bases. *Math. Comput.* **2001**, *70*, 863–872. [CrossRef]
10. Jiang, J.; Deng, Y. Strong pseudoprimes to the first 9 prime bases. *arXiv* **2012**, arXiv:1207.0063v1.
11. Sorenson, J.; Webster, J. Strong pseudoprimes to twelve prime bases. *arXiv* **2015**, arXiv:1509.00864v1.
12. Erdos, P.; Pomerance, C. On the number of false witnesses for a composite number. *Math. Comput.* **1986**, *46*, 259–279. [CrossRef]

13. Damgard, I.; Landrock, P.; Pomerance, C. Average case error estimates for the strong probable prime test. *Math. Comput.* **1993**, *61*, 177194. [CrossRef]
14. Ishmukhametov, S.; Mubarakov, B. On practical aspects of the Miller–Rabin primality test. *Lobachevskii J. Math.* **2013**, *34*, 304–312 [CrossRef]