Bollettino
Unione Matematica Italiana

R.G. Buschman, M.C. Wunderlich

# Sieves with generalized intervals.

*Bollettino dell'Unione Matematica Italiana, Serie 3*, Vol. **21** (1966), n.4, p. 362–367.

Zanichelli

<http://www.bdim.eu/item?id=BUMI_1966_3_21_4_362_0>

# Sieves with generalized intervals

R. G. Buschman and M. C. Wunderlich (Buffalo, N.Y., U.S.A.) (*)

**Summary.** - *The effects of alterxing the length of the sieving interval in a modification of the lucky number sieve are investigated. The asymptotic behavior of the $n^{th}$ term of such sieve generated sequences is considered for cases both with and without feedback.*

In recent papers of W. E. Briggs [1] and M. C. Wunderlich [2] sieve generated sequences are considered in which the sieving interval is of length $a_n$, the new element retained at the previous sieving stage. We will consider sieves with intervals of length $\mu_n$, where $\mu_n > 1$. (If $\mu_n \leq 1$, the sequence terminates). To define the sieve process let $A^{(1)} = \{a_k^{(1)}\}$, $a_k^{(1)} = k + 1$, and let $A^{(n+1)}$ be derived from $A^{(n)}$ by deleting the $r_{n,k}$ — *th* element of the interval $I_k^{(n)} = (n + (k-1)\mu_n,\ n + k\mu_n]$ where $1 \leq r_{n,k} \leq \mu_n$. The ultimate sequence $A = \{a_k\}$ is defined by $a_k = a_k^{(k)}$. In particular, for certain subclasses of these sequences, we will obtain a lower bound on $a_k$ and then an asymptotic estimate for $a_k$.

Let $f_n(x)$ denote the number of elements not exceeding $x$ which are sieved out of $A^{(n)}$ to produce $A^{(n+1)}$. Then $f_n(x) = R_n(x) - R_{n+1}(x)$ where $R_n(x)$ denotes the number of elements of $A^{(n)}$ which do not exceed $x$. If $x \in I_k^{(n)}$ we have two cases, depending upon whether $n + r_{n,k} + (k-1)\mu_k$ exceeds $x$ or does not exceed $x$; that is

$$f_n(x) = k - 1 \quad \text{or} \quad = k \quad \text{if} \quad n + (k-1)\mu_n < R_n(x) \leq n + k\mu_n.$$

This yields the formula

$$f_n(x) = [(R_n(x) - n)\ \mu_n] + \varepsilon_n$$

where $\varepsilon_n = 0$ or $= 1$. Following the methods of Briggs, if we now set

$$\sigma_n = \prod_{k=1}^{n} (1 - 1/\mu_k)$$

we have by analogous steps

$$R_{n+1}(x) = \sigma_n([x] - 1) + E_n(x), \qquad E_n(x) = \sum_{k=1}^{n} E_{k,n}(x),$$

where

$$E_{k,n}(x) = (\sigma_n/\sigma_k)(\{R_k(x) - k)/\mu_k\} + k/\mu_k - \varepsilon_k).$$

(*) The research conducted by the second author was supported in part by N.S.F. contract GP-1823, University of Colorado.

In order to estimate $E_n(x)$, we note that $\sigma_n/\sigma_k \leq 1$ and since

$$-1 \leq E_{k,n}(x) < 1 + k/\mu_k,$$

after summing we obtain

$$-n \leq E_n(x) < n + \sum_{k=1}^{n} k/\mu_k.$$

Since $R_{n+1}(a_n + 1) = n = \sigma_n a_n + E_n(a_n + 1) > \sigma_n a_n - n$, we have $\sigma_n a_n < 2n$. From $1/\sigma_k - 1/\sigma_{k-1} = (\sigma_k \mu_k)^{-1}$ it follows that

$$1/\sigma_n = 1 + \sum_{k=1}^{n} (\sigma_k \mu_k)^{-1}.$$

This can be estimated if there is a reasonable connection (feed-back) between $\mu_n$ and $a_n$. Let $\mu_n = \lambda_n a_n$ so that

$$1/\sigma_n = 1 + \sum_{k=1}^{n} (\lambda_n \sigma_n a_n)^{-1} > 1 + \frac{1}{2} \sum_{k=1}^{n} (k\lambda_k)^{-1}.$$

If there is no connection (no feedback), then since

$$a_k > (1 + 1/\mu_1)k \quad \text{we have} \quad (k\lambda_k)^{-1} = a_k/(k\mu_k) \geq (1 + 1/\mu_1)k$$

Hence

$$1/\sigma_n > 1 + \frac{1}{2}(1 + 1/\mu_1) \sum_{k=1}^{n} 1/\mu_k$$

and $1/\sigma_n$ can be estimated directly from the $\mu_k$'s.

In order to obtain a lower estimate on $a_n$ we note that for all $m$

$$n \leq R_{m+1}(a_n + 1) = \sigma_m a_n + E_m(a_n + 1) < \sigma_m a_n + m - 1 + \sum_{k=1}^{m} k/\mu_k.$$

This yields

$$a_n > (1/\sigma_m)(n + 1 - m - \sum_{k=1}^{m} k/\mu_k).$$

For the lower estimate to be useful we make

ASSUMPTION 1. - $\mu_n/n \geq \alpha > 0$.

(For the feedback case this is realized if $\mu_n = \lambda_n a_n$ where $\lambda_n \geq \lambda > 0$, for then $\mu_n \geq \lambda a_n \geq \lambda(1 + 1/\mu_1)n$). From Assumption 1 we have

$$a_n > (1/\sigma_m)(n + 1 - (1 + 1/\alpha)m)$$

which is positive if $n = (j+1)[1+1/\alpha]m$ $(j \geq \alpha)$, since

$$n+1-(1+1/\alpha)m > (j+1)(1/\alpha)m - (1+1/\alpha)m = (j/\alpha - 1)m.$$

This results in the estimate

$$a_n > (1/\sigma_m)(j/\alpha - 1)m = cn\,\sigma_{an} \quad \text{with} \quad a = ((j+1)[1+1/\alpha])^{-1}.$$

Some special examples of interest are

1. $\mu_k = \lambda_k a_k$ $(\lambda \leq \lambda_k \leq \Lambda)$, $a_n > cn \log n$;
2. $\mu_k = (\log k) a_k$, $a_n > cn \log\log n$;
3. $\mu_k = k a_k$, $a_n > cn$;
4. $\mu_k = \lambda k \log k$, $a_n > cn \log\log n$;
5. $\mu_k = \lambda k$, $a_n > cn \log n$.

(In each of these examples we assume that $\mu_1, \mu_2, \ldots, \mu_i$ are separately defined if necessary to insure that $\mu_k > 1$ for all $k$ and the given formulas are used only for large $k$).

Asymptotic estimates can next be generated analogous to WUNDERLICH [2]. We split the sum $E_n(a_n+1)$ into three parts (assuming $q(n)$ tends to infinity).

$$E_n(a_n+1) = \Sigma_0 + \Sigma_1 + \Sigma_2,$$

where $\Sigma_0$ extends over all $q(n) < k \leq n$ with $f_k(a_n) = 0$, $\Sigma_1$ over $q(n) < k \leq n$ with $f_k(a_n) = 1$, and $\Sigma_2$ over $1 \leq k \leq q(n)$. Since $E_{k,n}(a_n+1) < 1 + 1/\alpha$ by Assumption 1, $\Sigma_2 = \mathcal{O}(q(n))$. If $f_k(a_n)$ equals 0 or 1 then from the formula for $f_k(a_n)$ we can simplify the $E_{k,n}(a_n+1)$ terms obtaining, respectively,

$$E_{k,n}(a_n+1) = (\sigma_n/\sigma_k)((R_k(a_n+1))/\mu_k).$$

and

$$= (\sigma_n/\sigma_k)((R_k(a_n+1))/\mu_k - 1)$$

Since for $f_k(a_n) = 0$, $R_k(a_n+1) < 2n$ and for $f_k(a_n) = 1$, $R_k(a_n+1) < 3n$;

$$\Sigma_0 < n\mathcal{O}(S(n)), \qquad \Sigma_1 < \mathcal{O}(S(n)) - \sum_{k>q(n)}{}_1 (\sigma_n/\sigma_k),$$

where

$$S(n) = \sum_{k>q(n)}^{n} 1/\mu_k.$$

Next we consider the sum involving $\sigma_n/\sigma_k$,

$$1 \geq \sigma_n/\sigma_k = \prod_{k>q(n)}^{n} (1 - 1/\mu_k)$$

$$= \prod_{k>q(n)}^{n} ((1 - 1/\mu_k)e^{1/\mu_k}) \exp\left(- \sum_{k>q(n)}^{n} 1/\mu_k\right).$$

The first factor is a partial product of a convergent product, hence tends to 1. so that

$$\prod_{k>q(n)}^{n} ((1 - 1/\mu_k)e^{1/\mu_k}) = \exp\left(\sum_{k>q(n)}^{n} (\log (1 - 1/\mu_k) + 1/\mu_k)\right)$$

$$= \exp\left(\sum_{k>q(n)}^{n} \mathcal{O}(1/\mu_k^2)\right)$$

$$= \exp (\mathcal{O}(1/q(n)))$$

$$= 1 + \mathcal{O}(1/q(n)).$$

The second factor can also be estimated since

$$\exp\left(- \sum_{k>q(n)}^{n} 1/\mu_k\right) = \exp (- S(n)) \geq 1 - S(n).$$

Putting these together we obtain

$$1 \geq \sigma_n/\sigma_k \geq 1 - \mathcal{O}(S(n)) + \mathcal{O}(1/q(n))$$

so that

$$\Sigma_1 (1) \geq \sum_{k>q(n)}^{n}{}_1 \sigma_n/\sigma_k \geq \Sigma_1 (1) + n\mathcal{O}(S(n)) + \mathcal{O}(n/q(n)).$$

We now let $l(n)$ denote the number of $k$ such that $f_k(a_n) = 1$, so that

$$- \sum_{k>q(n)}^{n} \sigma_n/\sigma_k = - l(n) + \mathcal{O}(q(n)) + n\mathcal{O}(S(n)) + \mathcal{O}(n/q(n)).$$

This yields finally

$$E_n(a_n + 1) = - l(n) + \mathcal{O}(q(n)) + n\mathcal{O}(S(n)) + \mathcal{O}(n/q(n)).$$

We note that if $q(n) = o(n)$, $n/q(n) = o(n)$, $S(n) = o(1)$, then

$$E_n(a_n + 1) = - l(n) + o(n).$$

From this the case of feedback, $\mu_n = \lambda_n a_n$, results in

$$\sigma_n a_n \sim n + l(n) \quad \text{or} \quad \sigma_n \mu_n \sim \lambda_n (n + l(n)).$$

If we set $k/(k+l(k)) = d(k)$, then

$$1/\sigma_n \backsim 1 + \sum_{k=1}^{n} d(k)\,(k\lambda_k).$$

Also, however,

$$1/\sigma_n \backsim a_n/(n+l(n)) = a_n d(n)/n,$$

so that

$$a_n \backsim (n/d(n))(1 + \sum_{k=1}^{n} d(k)/k\lambda_k).$$

This would lead to formulas analogous to WUNDERLICH [2] in that

$$a_n \backsim n \log n \text{ iff} \left(1 + \sum_{k=1}^{n} d(k)/(k\lambda_k)\right) \backsim d(n) \log n$$

or in a more symmetric form

$$a_n \backsim (n \log n)/\lambda_n \text{ iff } (1 + \sum_{k=1}^{n} d(k)(k\lambda_k)) \backsim (d(n) \log n)/\lambda_n .$$

If we now consider the subclass of sequences which satisfy

ASSUMPTION 2. - $\mu_n > cnL(n)$, where $L(n)$ is an $m$-fold iterated logarithm ($m \geq 2$).

(This assumption is satisfied in the feedback case if $\mu_n = \lambda_n a_n$ where $\lambda_n \geq \lambda > 0$ and $a_n > cnL(n)$). From Assumption 2

$$S(n) = \sum_{k>q(n)}^{n} 1/\mu_n < \sum_{k>q(n)}^{n} (cnL(n))^{-1}$$

However, $q(n)$ is the largest $q(n)$ such that $q(n) + \mu_{q(n)} < c_1 n$ so that $c_2 n/L(n) < q(n) < c_3 n/L(n)$ and

$$S(n) < \mathcal{O}\left(\int_{q(n)}^{n} (xL(x))^{-1} dx\right) = \mathcal{O}(\log n/L(n) - \log q(n)/L(q(n))).$$

Since $\log q(n) = \log n - \log L(n) + o(\log L(n))$ and $L(q(n)) \backsim L(n)$,

$$\begin{aligned} S(n) &< \mathcal{O}(\log n/L(n) - (\log n/L(n))(1 - \log L(n)/\log n)) \\ &= \mathcal{O}(\log L(n)/L(n)) = o(1). \end{aligned}$$

Thus $q(n) = o(n)$, $n/q(n) = o(n)$, $S(n) = o(1)$ so that the results hold. We summarize this into the

THEOREM. – If $a_n > cnL(n)$, where $L(n)$ is an $m$-fold iterated logarithm $m \geq 2$, and $\mu_n = \lambda_n a_n$, $\lambda_n \geq \lambda > 0$, then

$$a_n \sim n \log n \text{ iff } (1 + \sum_{k=1}^{n} d(k)/(k\lambda_k)) \sim d(n) \log n.$$

Also, this can be written in the form

$$a_n \sim (n \log n)/\lambda_n \text{ iff } (1 + \sum_{k=1}^{n} d(k)/(k\lambda_k) \sim (d(n) \log n)/\lambda_n.$$

The special examples 1 and 2 satisfy the assumption. We note in particular that if $\lambda_n = \lambda$ then from example 1

$$a_n \sim (n \log n)/\lambda \text{ iff } 1 + (1/\lambda) \sum_{k=1}^{n} d(k)/k \sim (d(n) \log n)/\lambda.$$

Since $\frac{1}{2} \leq d(n) \leq 1$, in example 2 we have

$$A_1 + \frac{1}{2} \log \log n \leq A + \sum_{k=j}^{n} d(k)/(k\lambda_k) \leq A_2 + \log \log n,$$

but $\frac{1}{2} \leq (d(n) \log n)/\lambda_n \leq 1$, which yields the conclusion that $a_n \sim n$ and $a_n \sim n \log n$ do not hold.

For the cases in which $\mu_n$ is explicitely given (no feedback) $1/\sigma_n$ can be computed directly so that for examples 4 and 5 we have, repectively,

$$1/\sigma_n \sim c(\log n)^{1/\lambda} \quad \text{and} \quad 1/\sigma_n \sim cn^{1/\lambda}.$$

Hence from $a_n \sim (1/\sigma_n)(n + l(n))$ we have, respectively,

$$a_n \sim cn(\log n)^{1/\lambda}(1 + l(n)/n) \quad \text{and} \quad a_n \sim cn^{1+1/\lambda}(1 + l(n)/n).$$

It is interesting to note the different role played here by the scalar multiplier $\lambda$ from that in the case of feedback.

## REFERENCES

[1] W. E. BRIGGS, *Prime-like sequences generated by a sieve process*, « Duke Math. J. », 30 (1963), 297-312.

[2] M. C. WUNDERLICH, *Sieve-generated sequences*, « Canad. J. Math. », Vol. 18, pp. 291-299.

*Pervenuta alla Segreteria dell'*U.M.I.
*il 22 settembre 1965*