On constructing sequences of radius k using finite geometries

Joshua W. Gilkerson[1], Jerzy W. Jaromczyk[1**] and Zbigniew Lonc[2]

[1] University of Kentucky, Lexington, KY, USA. <{jwgilk0,jurek}@cs.uky.edu>,
[2] Technical University, Warsaw, Poland. <zlonc@mini.pw.edu.pl>

Dedicated to Godfried Toussaint on the occasion of his 60th birthday

Abstract. We say that a sequence $a_1, a_2, \dots a_m$ over n element set Σ has the *k-radius* property (or, simply, has radius k) if every pair of different elements in Σ occurs at least once within distance at most k; where the *distance* $\mathrm{d}(a_i, a_j) = |i - j|$. Recent paper [JL04] presented constructions of k-radius sequences for arbitrary values of k and n. This paper elaborates on sequences constructed with finite geometries. For example, sequence 142536475162731 of radius 2 is based on the Fano projective plane $P_2(7)$.

1 Introduction

Let $a_1, a_2, \dots, a_m$ be a sequence over an n-element set Σ, such as a set $[n]$ of n integers. We say that a sequence $a_1, a_2, \dots a_m$ has the *k-radius* property if every pair of different elements in Σ occurs at least once within distance at most k; *distance* $\mathrm{d}(a_i, a_j) = |i - j|$. In other words, each pair of objects will appear at least once inside a window of size $k + 1$ that slides along the sequence.

The problem of finding short k-radius sequences originated in the context of computing a two-argument function for all pairs of n large objects, such as images or matrices or bitmaps [GJ02]. The restriction is that the limited memory size, which is a constant $k + 1$, prohibits us from storing all the objects at the same time. The goal is to schedule the shortest possible sequence of *read* operations that will ensure that, for all pairs i, j, there is some point in time when both a_i and a_j will reside in memory. Such a scenario may also appear in processing a large quantity of huge objects located in a remote database when locally storing fetched data may be either impossible or impractical and where the limited bandwidth and time require efficient scheduling of the data requests. Additionally, k-radius sequences may be viewed as a method for efficient scheduling for the caching process [SChD02] for some computations. Clearly, for a given n and k sequences differ in the order of elements as well as in the length. It is possible to construct sequences that begin with a predefined sequence of elements, as in the following examples for $n = 26$ and k equal 3, 4 and 5, respectively:

** Corresponding author.This research was supported in part by KBRiN, KY-INBRE, KY NSF EPSCOR grants.

```
GODFRIED ABCHJKLMNPQS TUVWXYZGIB KOPAJRNTCE LQVFHMSWDJ
QXIOCUYFAT ZMOEGHPUZD KVRBLSXAGN WIBLYMRXHT KYESJUZNLD
OVCGPWFKQB GRTIVNHYDP IMCSURWZQA VJFXEWOSGJ LBUCZDTFNP
XQY

GODFRIED ABCHJKLMNPQS TUVWXYZGEH NUOFAKQWIR CMTYOBLSXF
JPVZDILUGJ RBVNPCAXGS KDTEZMAQYJ DWNEPHIVOX MZRSHQBTFY
CMUGQXLWBO JZCTYIKPSV LAE

GODFRIED ABCHJKLMNPQS TUVWXYZGEJ QSOFBCYNPR WIHXKOTZAL
MDUVFKEHGT BMRUJIVLWB XDNAQPGZCI SRFXHQYMLE ASDKYCVOUN
EPWFMBZCTG L
```

We are interested in the question: *How to construct short sequences with the* k-radius *property?*

Various methods have been presented in [JL04]. In this paper we focus on constructing k-radius sequences using finite geometries. For example, 142536475162731 for $n = 7$ of radius 2 is constructed using the Fano projective plane of order 2.

2 Review of known results, lower bounds

A related result, formulated in the context of the consecutive 1 property for data bases and asking for all pairs appearing consecutively in a sequence, was obtained by Ghosh [G75]; see also [LTT81]. In our terminology the sequence has 1-radius property.

Theorem 1. *(Ghosh 1975)*

$$f_1(n) = \begin{cases} \binom{n}{2} + 1, & \text{for } n \text{ odd;} \\ \binom{n}{2} + \frac{1}{2}n, & \text{for } n \text{ even.} \end{cases}$$

For n objects, the length $f_k(n)$ of the shortest sequence for n elements with the *k-radius* property is bounded from below as follows:

Theorem 2. *([JL04])*

$$f_k(n) \geq \left\lceil \frac{1}{k}\binom{n}{2} + \frac{k+1}{2} \right\rceil$$

or even slightly stronger

Theorem 3. *([JL04])*

$$f_k(n) \geq \left\lceil \frac{n-1}{2k} \right\rceil n + \sum_{j=1}^{k} \left(\left\lceil \frac{n+k-j}{2k} \right\rceil - \left\lceil \frac{n-1}{2k} \right\rceil \right)$$

For sequences of radius 2 we have the following:

Corollary 1.

$$f_2(n) \geq \begin{cases} \frac{1}{2}\binom{n}{2} + \frac{1}{4}n + 1, & n \equiv 0 \pmod 4 \\ \frac{1}{2}\binom{n}{2} + 2, & n \equiv 1 \pmod 4 \\ \frac{1}{2}\binom{n}{2} + \frac{3}{4}n, & n \equiv 2 \pmod 4 \\ \frac{1}{2}\binom{n}{2} + \frac{1}{2}n, & n \equiv 3 \pmod 4 \end{cases}$$

Specific lower bounds for $f_2(n)$ and the length of the best sequences we have constructed (not based on finite geometries) for small values of n are listed in the following table:

n	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
lower bounds	2	3	5	7	12	14	17	20	30	33	37	41	56	60	65	70	90
best found	2	3	5	7	12	14	20	24	31	34	44	50	58	66	73	83	94

The lengths of the best found are not necessarily optimal and sequences as we used non-exhaustive ad-hoc search algorithms. Example optimal sequences for $k = 2$ and n equal $2, 3, 4, 5$ and 6 are 12, 123, 12341, 1234512 and 123456124536, respectively.

3 Constructions based on Steiner systems and finite geometries

Before switching to finite geometries [D68,HP73], we will look at Steiner systems [GGL95,CR99]. Sequences with the k-radius can be constructed with Steiner systems. The construction applies, however, only to the values of n for which the Steiner systems exist.

Consider a Steiner system $S(M, m, 1)$, i.e., a collection $\mathcal{B}$ of m-element subsets of an M-element set $X = \{x_1, x_2, \ldots, x_M\}$ such that every pair of elements in X is contained in exactly one block in $\mathcal{B}$.

Let k be an integer such that $k+1 = m \cdot \lceil \frac{n}{M} \rceil$, $n \geq M$. We partition the set $[n] = \{1, 2, \ldots, n\}$ into M disjoint subsets of cardinality $\lfloor \frac{n}{M} \rfloor$ or $\lceil \frac{n}{M} \rceil$. Let these subsets be $A_1, A_2, \ldots, A_M$.

For every block $B \in \mathcal{B}$ we define a subset of $[n]$, $N_B = \bigcup_{x_j \in B} A_j$. Observe that each pair of elements $r, t \in [n]$ is included in one of the sets N_B. Indeed, let $r \in A_p$ and $t \in A_q$. Clearly the pair $\{x_p, x_q\} \subseteq B$, for some block $B \in \mathcal{B}$. Then $r \in A_p \subseteq N_B$ and $t \in A_q \subseteq N_B$, by the definition of N_B.

Let s_B be any sequence (permutation) of elements of N_B and let s be a concatenation of all sequences s_B. We claim that every pair of elements $\{r, t\}$ is within distance at most k in the sequence. Indeed, as we have already checked, $r, t \in N_B$ for some block $B \in \mathcal{B}$. The claim follows now because $|N_B| = \sum_{x_j \in B} |A_j| \leq m \cdot \lceil \frac{n}{M} \rceil = k + 1$.

Let us compute the length of s.

$$|s| = \sum_{B\in\mathcal{B}} |N_B| = \frac{|\mathcal{B}|\cdot n\cdot m}{M} = \frac{\binom{M}{2}}{\binom{m}{2}}\cdot\frac{n\cdot m}{M} = \frac{M(M-1)}{m(m-1)}\cdot\frac{nm}{M} = n\cdot\frac{M-1}{m-1}$$

By Fisher's inequality [GGl95,CR99] $M \leq |\mathcal{B}| = \frac{\binom{M}{2}}{\binom{m}{2}} = \frac{M(M-1)}{m(m-1)}$, so $M \geq m^2 - m + 1$ and consequently $\frac{M-1}{m-1} \geq \frac{M}{m} + 1 - \frac{1}{m}$, so

$$|s| \geq n\left(\frac{M}{m} + 1 - \frac{1}{m}\right) = \frac{nM}{m} + \left(1 - \frac{1}{m}\right)n \geq \frac{nM}{m} \geq \frac{n^2}{k+1}$$

as $k+1 = n\cdot\left\lceil\frac{n}{M}\right\rceil \geq \frac{nm}{M}$.

In the next section, we will show upper bound on $|s|$ for finite projective and affine planes, which are Steiner systems for specific values of m and M

3.1 Finite projective and affine planes

A finite projective plane is a finite set of points and lines that satisfies the axioms:

- Any two points are on exactly one line.
- Any two lines intersect in exactly one point.
- There are four points, not three of which are collinear.

Projective planes are characterized by their order m and they have m^2+m+1 points and (by duality) the same number of lines. Projective plane of order m is a $S(m^2+m+1, m+1, 1)$ Steiner system.

Finite *affine planes* of order m are obtained from projective planes of order m by removing any one line and its points. Affine plane of order is a $S(m^2, m, 1)$ Steiner system.

There do not exist finite planes of order m for all values of m. A sufficient condition is that m is a power of a prime number. Although it is conjectured that prime powers are only possible orders, the strongest result to date is Bruck-Ryser [D68] theorem that says that if m is a positive integer of the form $4p+1$ or $4p+2$ and m is not equal to the sum of two integer squares, then m is not the order of a finite plane. Either by virtue of Bruck-Ryser theorem or by intensive computer search, it is for example known that there are no finite planes of order 6 and 10.

The simplest projective plane consists of seven points and seven lines and its order is 2. Each point is on three lines, and each line contains three points. This particular projective plane is called the Fano plane, named for Gino Fano (1871-1952), the Italian geometer. If any of the lines is removed from the Fano plane, along with the points on that line, the resulting geometry is the affine plane of order 2. A typical drawing of Fano plane is shown in Figure 1.

For $M = |\mathcal{B}|$, i.e, for quadratic Steiner systems, by virtue of the inequalities from the previous section, we have

$$|s| = m\cdot\frac{M-1}{m-1} = \frac{nM}{m} + \left(1 - \frac{1}{m}\right)n \leq \frac{n(n+M-1)}{k+1} + n \tag{1}$$

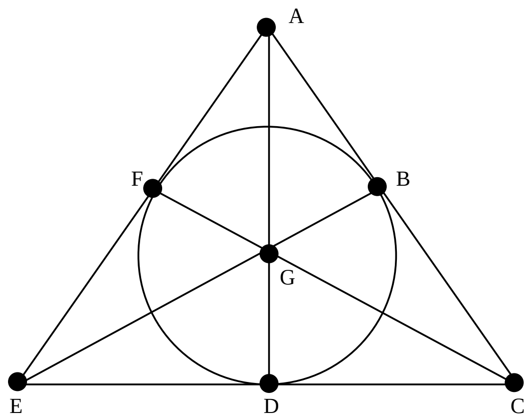


Fig. 1. Fano plane: 7 lines and 7 points with 3 points per line and each pair of lines intersecting in exactly one point

because $k+1 = \left\lceil \frac{n}{M} \right\rceil \cdot m \leq \frac{n+M-1}{M} \cdot m$.

If n is a multiple of $M = m^2 - m + 1$ then we have a stronger inequality

$$|s| = \frac{nM}{m} + \left(1 - \frac{1}{m}\right) n = \frac{n^2}{k+1} + \left(1 - \frac{1}{m}\right) n \leq \frac{n^2}{k+1} + n \tag{2}$$

In particular, we have inequalities (1) and (2) if a projective plane of the order $m-1$ exists (i.e. $M = m^2 - m + 1$), for $k+1 = \left\lceil \frac{n}{m^2-m+1} \right\rceil \cdot m$.

Using an affine plane of the order m (which is a Steiner system as well, with $M = m^2$), we get the following inequalities.

$$|s| = n \cdot \frac{M-1}{m-1} = n \cdot \frac{m^2-1}{m-1} = n(m+1) \leq n\left(\frac{n+M-1}{k+1} + 1\right) \tag{3}$$
$$= \frac{n(n+M-1)}{k+1} + n$$

because $k \leq \frac{n+M-1}{M} \cdot m = \frac{n+M-1}{m}$.

When n is divisible by $M = m^2$ then we get a precise formula

$$|s| = n(m+1) = \frac{n^2}{k+1} + n \tag{4}$$

because $k+1 = \left\lceil \frac{n}{M} \right\rceil \cdot m = \frac{n}{m^2} \cdot m = \frac{n}{m}$.

So inequality (3) and equation (4) are satisfied in particular for those values of n and k for which an affine plane of order m (i.e. $M = m^2$) exists and $k+1 = \left\lceil \frac{n}{m^2} \right\rceil \cdot m$. An important special case is n divisible by m and $k+1 = \frac{n}{m}$.

3.2 Constructions

The above results for finite geometries suggest a simple construction for k-radius sequences if a number of elements and k allow for the existence of finite geometry of the adequate order. Take a finite geometry of order m. Divide n elements into the number of groups equal to the number of points in the geometry; m^2 for affine and $m^2 + m + 1$ for projective plane, assuming that n is a multiple of the number of points. Than assign each group to a point and list all the lines in any order substituting points on each line with their corresponding group of elements. Clearly, the length of the sequence is equal to the product of the number of lines, times the number of points per line, times the number of elements in each group. The radius of the sequence is one less than the number of points per line (which is $m + 1$) multiplied by the number of elements associated with each point.

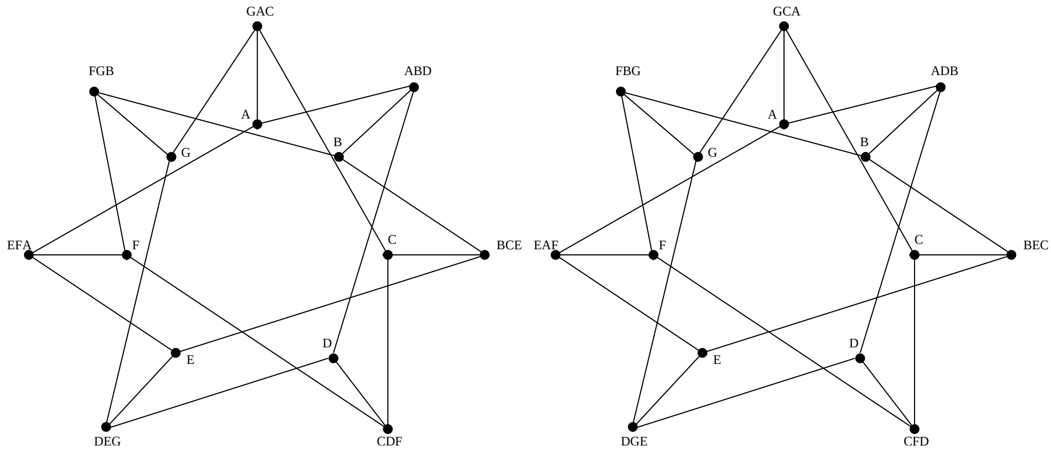
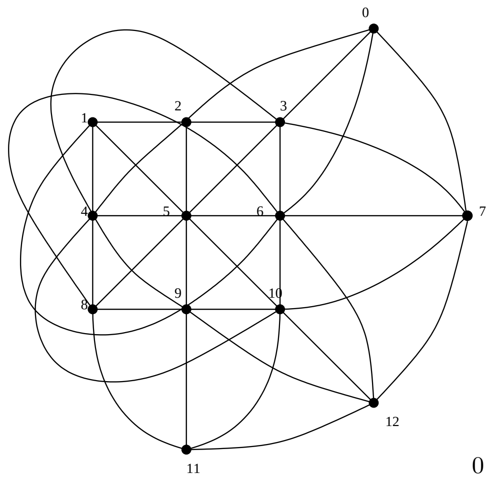


Fig. 2. Graphs illustrating lines and their points based on the Fano plan e. Vertices labeled with three letters correspond to the seven lines of the Fano plane and are connected in the graph with the vertices corresponding to the points that these lines contain. The labels for the lines in the left and right picture differ only by the order of the letters.

A number of examples for projective planes of order 2 and order 3 will illustrate this construction. Let us start with the Fano plane. Figure 2 show the vertices corresponding to lines in Fano plane with edges connecting with the corresponding points. As mentioned above, a list of the lines and the points on them gives us a sequence of length $3 \cdot 7$ where each letter can be replaced with a group of elements. If we read the points from Figure 2 (left) in the clockwise direction starting from the top, we obtain $GAC\ ABD\ BCE\ CDF\ DEG\ EFA\ FGB$.



0 1 2 3	0 4 5 6	0 8 9 12	0 7 10 11	1 4 7 8	1 6 9 11
1 5 10 12	3 4 9 10	2 4 11 12	2 5 7 9	3 6 7 12	3 5 8 11
2 6 8 10					

Fig. 3. Projective plane of order 3

Since the order of the points on a line is not important for the construction, with a different permutation of them, as presented in Figure 2 (right), we obtain GCA ADB BEC CFD DGE EAF FBG. Clearly, without breaking the 2-radius property, one element from each of 6 pairs of the adjacent identical elements can be removed, resulting in a sequence of the length 15 rather than 21. In the above sequence element G does not appear back to back (it does if we look at the sequence as circular) so it cannot be replaced with one occurance. Note from the lower bounds that for $k = 2$ and $n = 7$ the sequence is not shorter than 14 and therefore the following sequence $CEGADBEGFADCFB$ is optimal.

Similarly, we can construct k-radius sequences for specific values of k and n based on the projective plane of order 3; see Figure 3.

Based on this list, we can construct, for example, a 7-radius sequence for $n = 26$. Figure 4 shows this sequence of length 104, the number of times each pair occurs in distance less than $k + 1$ and the closest distance for each pair of distinct elements.

A	B	C	D	E	F	G	H
A	B	I	J	K	L	M	N
A	B	Q	R	S	T	Y	Z
A	B	O	P	U	V	W	X
C	D	I	J	O	P	Q	R
C	D	M	N	S	T	W	X
C	D	K	L	U	V	Y	Z
G	H	I	J	S	T	U	V
E	F	I	J	W	X	Y	Z
E	F	K	L	O	P	S	T
G	H	M	N	O	P	Y	Z
G	H	K	L	Q	R	W	X
E	F	M	N	Q	R	U	V

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	
	1	2	3	4	3	2	1	2	3	4	3	2	1	2	3	2	3	4	3	4	5	6	7	2	1	A
		1	2	3	4	3	2	1	2	3	4	3	2	1	2	1	2	3	4	3	4	5	6	3	2	B
			1	2	3	4	5	2	3	2	3	2	3	4	3	2	1	4	3	4	3	2	1	6	7	C
				1	2	3	4	1	2	1	2	1	2	3	4	3	2	3	4	3	4	3	2	5	6	D
					1	2	3	2	3	2	3	2	3	4	5	4	3	4	3	2	1	2	1	2	1	E
						1	2	1	2	1	2	1	2	3	4	3	4	5	4	3	2	3	2	3	2	F
							1	2	3	2	3	2	3	4	3	4	5	2	1	4	3	6	7	2	1	G
								1	2	1	2	1	2	3	4	3	4	3	2	5	4	5	6	3	2	H
									1	2	3	4	5	2	3	4	5	2	3	4	3	2	3	4	3	I
										1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	J
											1	2	3	2	3	2	3	4	5	2	3	4	3	4	3	K
												1	2	1	2	1	2	3	4	1	2	3	4	3	4	L
													1	2	3	2	3	2	3	4	5	4	3	4	5	M
														1	2	1	2	1	2	3	4	3	4	3	4	N
															1	2	3	2	3	2	3	4	5	2	3	O
																1	2	1	2	1	2	3	4	1	2	P
																	1	2	3	2	3	2	3	4	5	Q
																		1	2	1	2	1	2	3	4	R
																			1	2	3	2	3	2	3	S
																				1	2	1	2	1	2	T
																					1	2	3	2	3	U
																						1	2	1	2	V
																							1	2	3	W
																								1	2	X
																									1	Y
																										Z

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	
	7	2	2	2	2	2	2	2	2	2	2	2	1	1	2	2	2	2	1	1	1	1	2	2	A
		3	2	2	2	2	2	2	2	2	2	2	1	1	2	2	2	2	1	1	1	1	2	2	B
			6	1	1	1	1	2	2	1	1	2	2	3	3	2	2	2	2	2	2	3	3	1	1	C
				1	1	2	1	3	2	1	1	2	2	3	3	2	2	2	2	2	2	3	3	1	1	D
					5	1	3	4	4	2	2	1	1	1	1	2	2	2	2	2	2	3	3	2	2	E
						2	1	4	4	3	2	1	1	1	1	2	2	2	2	2	2	3	3	2	2	F
							5	2	2	4	4	2	2	3	3	1	1	2	2	2	2	1	1	3	3	G
								2	2	4	4	3	2	3	3	1	1	2	2	2	2	1	1	3	3	H
									5	1	2	1	1	1	2	1	1	2	2	4	4	2	2	2	2	I
										2	1	2	1	1	1	2	1	2	2	4	4	2	2	2	2	J
											4	1	3	2	2	2	2	2	2	1	1	3	3	3	3	K
												3	1	2	2	2	2	3	2	1	1	3	3	3	3	L
													4	3	3	4	4	3	3	1	1	2	2	1	1	M
														3	3	4	4	3	3	1	1	2	2	2	1	N
															5	1	2	3	3	1	2	2	3	3	3	O
																2	1	4	3	1	1	2	2	3	3	P
																	5	2	2	1	1	2	3	2	2	Q
																		2	2	1	1	3	2	2	2	R
																			4	1	2	1	1	2	3	S
																				3	1	2	1	3	2	T
																					4	3	4	2	3	U
																						3	3	3	2	V
																							4	1	3	W
																								2	1	X
																									4	Y
																										Z

Fig. 4. Sequence based on $P_3(13)$, the number of occurrences at distance less than 8 for each pair, and the smallest distance for each pair. Out of 325 pairs, 62 are at the minimal distance 1, 98 at distance 2, 88 at distance 3, 52 at distance 4, 16 at distance 5, 6 at distance 6 and 3 at distance 7.

The below table shows lengths of k-radius sequences for various values of n and k constructed with an *ad hoc* search algorithm, constructed with the projective plane of order 3, and the corresponding lower bound based on Theorem 2.

k	3	7	11	15	19	23	27	31	35	39
n	13	26	39	52	65	78	91	104	117	130
search	39	81	121	165	201	245	287	328	368	407
$P_k(n)$	52	104	156	208	260	312	364	416	468	520
$f_k(n)$	29	56	83	110	137	164	191	218	245	272

4 Conclusions

Sequences of radius k can be constructed in a number of ways, including methods based on radius 2 and radius 3 sequences, as presented in [JL04]. The construction based on the Steiner systems, in particular on finite geometries, as presented on this paper, are characterized by their high regularity as the proximity between elements occurs in blocks of a fixed size. By permuting the blocks and elements in each block and removing adjacent identical elements, we can shorten the sequence, as we demonstrated in the Fano projective plane case. However, for some applications, e.g., where caching based on blocks of the same length is desired, regularity offered by the finite geometries may be beneficial. A natural extension of our construction is to consider sequences where each combination of r elements, $r \geq 2$, occurs within a distance of not more than k.

Bibliography

[CR99] Colbourn, Ch., J., Rosa, A., Triple Systems, Clarendon Press - Oxford, 1999.

[D68] Dembowski, P., Finite Geometries, Springer-Verlag, Berlin, 1968.

[HP73] Hughes, D.R., Piper, F.C., Projective Planes, Springer-Verlag, New York, 1973

[GJ02] Gilkerson, J. W., Jaromczyk, J. W., Restoring the order of images in a sequence of MRI slices, Manuscript, 2002.

[G75] Gosh, S.P., Consecutive storage of relevant records with redundancy, *Communications of the ACM* 18(8):464-471, 1975.

[GGL95] Graham, R. L., Grötschel, M., Lovász, L., Handbook of Combinatorics, The MIT Press - North Holland, 1995.

[GKP89] Graham, R. L., Knuth, D. E., Patashnik, O., Concrete Mathematics, Addison-Wesley, 1989.

[JL04] Jaromczyk, J. W., Lonc, Z., Sequences of radius k: how to fetch many huge objects into small memory for pairwise computations. *to appear in* the Proceedings of ISAAC 2004, Hong Kong, December 2004.

[LTT81] Lonc Z., Traczyk T. Truszczyński, M., Optimal f-graphs for the family of all k-subsets of an n-set. Data base file organization (Warsaw, 1981), 247–270, Notes Rep. Comput. Sci. Appl. Math., 6, Academic Press, New York, 1983.

[SChD02] Sen, S., Chatterjee, S., Dumir, N., Towards a theory of cache-efficient algorithms, Journal of the ACM, 49(6):828-858, 2002.