A COMPARISON OF TWO INVARIANTS FOR STEINER TRIPLE SYSTEMS: FRAGMENTS AND TRAINS

D.R. Stinson

ABSTRACT

We desire invariants for Steiner triple systems that are fast, effective, and require little space. Two invariants are described and compared: fragment vectors and indegree sequences of trains. Both invariants can be computed in time $0(v^3)$, and require space $0(v)$, where v is the order of the Steiner triple system. Experimental evidence suggests that the invariant based on trains is more effective, but it requires about five times longer to compute. However, both methods appear very well suited to practical applications.

1. Introduction.

A *Steiner triple system* (or STS) of order v is a pair (X,B), where X is a set of v elements called *points*, and B is a set of 3-subsets of X (called *blocks*) such that every pair of points occurs in a unique block. It follows that every point occurs in $r = (v-1)/2$ blocks, and there are $b = v(v-1)/6$ blocks. Hence $v \equiv 1$ or 3 mod 6. This necessary condition for existence is also sufficient, as was demonstrated by Kirkman [5] in 1847.

Let $D_i = (X_i, B_i)$ be STS of order v, for $i = 1,2$. We say that D_1 and D_2 are isomorphic if there is a bijection ϕ: $X_1 \rightarrow X_2$ such that $\{x,y,z\} \in B_1$ if and only if $\{\phi(x),\phi(y),\phi(z)\} \in B_2$. The number of pairwise non-isomorphic STS of order is denoted by $N(v)$. It is known that $N(v) = 1$ for $V = 1,3,7,$ or 9; $N(13) = 2$; and $N(15) = 80$. Beyond this point $N(v)$ grows extremely rapidly: $N(19) \geq 284407$, and, in general $N(v) = v^{v^2(1/6+0(1))}$. (References for all these results can be found in [6].)

Given two STS of order v, how difficult is it to test them for isomorphism? An algorithm which works in time $0(v^{\log v})$ is presented in Miller [7]. No polynomial-time algorithm is known, though Miller's algorithm, in practice, appears to run in time $0(v^4 \log v)$, on average. An implementation of this algorithm, due to the author, is described in [9]. We found that to test isomorphism of typical STS of order 15 required 1.6 seconds (on an AMDAHL 470/V8 computer). We want a faster test, one that will always run in polynomial time. Such tests exist, though they do not always succeed.

An *invariant* is a mapping f, defined on the set of all STS, such that $f(D_1) = f(D_2)$ if D_1 and D_2 are isomorphic. The image $f(D)$ of an STS D is called a *form*. Notice that two STS may conceivably have the same forms, yet not be isomorphic. So, if two STS have the same forms, we must use some other invariant to prove that they are non-isomorphic, or alternatively, attempt to find an isomorphism.

We will evaluate invariants by three criteria: effectiveness, time and space.

The effectiveness of an invariant is the tendency that non-isomorphic STS is mapped to different forms. In [8], a measure for effectiveness was proposed. Let f be an invariant, let v be a positive integer, and let $S(v)$ denote the set of all STS of order v. The *sensitivity* of f is the function s where $s(v) = |\{f(B):B \in S(v)\}|/|S(v)|$. (The number of forms divided by the number of non-isomorphic STS.) Clearly $s(v)$ is between 0 and 1. We desire $s(v)$ to be close to 1; if $s(v)$ is unity, we say that f is ***complete*** at order v.

The time of an invariant will be regarded as a function of the order v of the STS being tested.

The space of an invariant is the amount of computer memory required to store a form. Throughout this paper, we will be informal and assume that all integers involved will fit into a single word of memory. (Certainly this will be true in any practical application of the methods we describe here.)

Clearly, we wish to maximize sensitivity, while minimizing time and space. It is currently unknown if there exists a complete invariant that requires only polynomial time and space.

In this paper we discuss two invariants, trains and cycle structure, which require time and space $0(v^3)$, but which are not complete. "Compressed" invariants based on these require only space $0(v)$. In section 4, we compare these two invariants, based mostly on empirical evidence.

2. Cycle structure and fragment vectors.

Cycle structure has been independently discovered by many researchers; Cole [2] and Cummings [3] appear to be the first.

Let $(X,\mathcal{B})$ be an STS of order v, let x,y be any two distinct points, and let $B = \{x,y,z\}$ be the block containing x and y. Define $G_{x,y} = \{\{a,b\}:\{a,b,x\} \in \mathcal{B}\backslash B\} \cup \{\{a,b\}:\{a,b,y\} \in \mathcal{B}\backslash B\}$. $G_{x,y}$ is a 2-regular graph on vertex set $X\backslash\{x,y,z\}$. $G_{x,y}$ consists of even cycles of length at least four, and hence can be represented as a partition of $v-3$ into even parts, each at least four. Ordered, this partition is called a *cycle list*. The collection of all $\binom{v}{2}$ cycle lists, ordered lexicographically, is called *cycle structure*. It is not difficult to see that one can determine cycle structure in both time and space proportional to v^3.

We do not object to time $O(v^3)$, but space $O(v^3)$ will prove intractible in any large-scale applications. Gibbons [4] has suggested a way of "compressing" the cycle structure by considering only cycles of length 4 (in the graphs $G_{x,y}$). For $x,y \in X$, let $a_{x,y}$ denote the number of four-cycles in $G_{x,y}$. For $x \in X$ let $f_x = \sum_{y \neq x} a_{x,y}$. The list $(f_x\colon\ x \in X)$, ordered, is called the *fragment vector* of the STS.

Note that we do not have to determine all the graphs $G_{x,y}$ in order to find the fragment vectors. A *fragment* is a set of four blocks of the form $\{u,v,w\}$, $\{u,x,y\}$, $\{v,x,z\}$, $\{w,y,z\}$. A fragment gives rise to a four-cycle in $G_{u,z}$, $G_{v,y}$ and $G_{w,x}$. We can determine the fragment vector simply by finding all fragments, and each time one is found, incrementing f_i by one for each of the six points i in the fragment. This method still requires time proportional to v^3, but is considerably quicker than determining the complete cycle structure. Note that to store a fragment vector requires space proportional to v. This is a considerable saving in space over the complete cycle structure.

How big are the numbers f_x in a fragment vector? It is not difficult to see that $f_x \leq (v-1)(v-3)/4$, with equality occurring if and only if $g_{x,y}$ is a union of 4-cycles, for all y. Also, $f_x = (v-1)(v-3)/4$ for all x if and only if the STS is a projective space $PG(n,2)$.

In [4], Gibbons lists the fragment vectors for all 80 STS of order 15. Since they are all different, fragment vectors are a complete invariant at $v = 15$.

3. Trains and indegree lists.

Trains were first described by White [11], and then largely ignored. Recently, Colbourn *et al* [1] discussed the use of trains as an invariant.

Let (X,B) be an STS of order v. For any two distinct points x,y, define other$(x,y) = z$, where $\{x,y,z\} \in B$. A *train* is a directed graph T whose vertices are the $\binom{v}{3}$ 3-subsets of X. T is regular of outdegree 1; the edge leaving $\{x,y,z\}$ is directed to {other(x,y), other(x,z), other(y,z)}.

It is not difficult to see that one can calculate a train in time and space proportional to v^3. As was the case in cycle structure, we desire a smaller invariant. Colbourn *et al* [1] define *compact trains:* a compact train is a set of triples (i,j,k); a triple (i,j,k) means that there are k components of T which have i vertices, j of which have indegree zero. Compact trains are certainly more compact that trains, but it is not clear how much more compact they are. For example, one STS of order 15 (#23) requires 29 triples for its compact train. Further, compact trains fail to distinguish two of the eighty STS of order 15 (#6 and #7).

We would like to propose an invariant based on the indegree sequence of trains.

Lemma 3.1. *No vertex in a train has indegree exceeding $v-2$. Further, any vertex of indegree $v-2$ is a block of the STS.*

Proof. Let $\{x,y,z\}$ be a 3-subset. There are $\frac{v-1}{2}$ pairs $\{a_i,b_i\}$ for which other $(a_i,b_i) = x$. $(1 \leq i \leq (v-1)/2)$. The point y is in one of these pairs; let $a_1 = y$. If $\{a_1,b_1,c\}$ is directed to $\{x,y,z\}$, then other$(y,b_1) = x$ and {other(y,c),other(b_1,c)} $= \{y,z\}$. Hence other$(y,c) = z$ and other$(b_1,c) = y$. Since $\{x,y,b_1\}$ and $\{c,y,b_1\}$ are blocks, we must have $x = c$. Then $z =$ other$(y,c) =$ other(x,y), so $\{x,y,z\}$ is a block. Hence $\{y,b_1,c\}$ is directed to $\{x,y,z\}$ if and only if $\{x,y,z\} = \{y,b_1,c\}$ is a block of the STS.

Now, consider any $\{a_i,b_i\}$, with $i > 1$. Suppose $\{a_i,b_i,c\}$ is directed to $\{x,y,z\}$. Either $c =$ other(a_i,y) or $c =$ other(b_i,y), so these are at most two choices for c.

It follows that $\{x,y,z\}$ has indegree at most $1 + 2(\frac{v-3}{2}) = v-2$, and equality is attained only when $\{x,y,z\}$ is a block. □

It is interesting to note that, in the projective space $PG(n,2)$, all blocks have indegree $v-2$, and all non-blocks have indegree 0.

Since the indegrees are at most $v-2$, we may form a vector $(a_i : 0 \leq i \leq v-2)$, where a_i is the number of vertices of indegree i. We refer to this as the ***indegree list*** of the train. The space required to store an indegree list is clearly proportional to v, so we have a "small" invariant. However, the time required is still proportional to v^3.

For STS of order 15, indegree lists distinguish all non-isomorphic designs. These lists are presented in Table 1. In fact, the ordered pair (a_0, a_2) is a complete invariant for STS of order 15. This is the most concise complete invariant known to the author. (The 80 STS are numbered as in [6]).

Table 1.
Indegree lists of STS of order 15

STS	indegree list													
1	420	0	0	0	0	0	0	0	0	0	0	0	0	35
2	392	0	0	0	24	4	0	0	4	24	0	0	0	7
3	354	0	48	0	12	6	0	16	6	12	0	0	0	1
4	348	8	20	20	22	8	12	4	2	10	0	0	0	1
5	348	16	0	32	20	10	16	0	4	8	0	0	0	1
6	336	0	16	44	36	10	0	12	0	0	0	0	0	1
7	336	0	0	48	52	18	0	0	0	0	0	0	0	1
8	300	40	44	22	15	16	10	4	3	1	0	0	0	0
9	288	48	39	34	24	10	8	1	1	2	0	0	0	0
10	290	40	41	46	18	6	10	1	1	2	0	0	0	0
11	261	61	55	49	18	6	2	2	0	1	0	0	0	0
12	287	44	48	42	12	3	13	3	0	3	0	0	0	0
13	289	48	48	16	26	17	8	0	1	2	0	0	0	0
14	296	48	48	6	21	22	6	4	1	3	0	0	0	0
15	269	48	64	38	22	11	0	2	1	0	0	0	0	0
16	329	0	84	0	0	7	0	28	7	0	0	0	0	0
17	275	36	72	24	40	7	0	0	1	0	0	0	0	0
18	273	48	50	48	24	9	0	2	1	0	0	0	0	0
19	230	86	78	42	12	3	4	0	0	0	0	0	0	0
20	254	59	66	54	15	3	1	3	0	0	0	0	0	0
21	245	68	81	39	15	0	1	6	0	0	0	0	0	0
22	236	77	81	45	9	0	4	3	0	0	0	0	0	0
23	219	107	67	46	9	2	5	0	0	0	0	0	0	0
24	218	116	59	42	10	5	5	0	0	0	0	0	0	0
25	224	104	68	35	15	4	5	0	0	0	0	0	0	0
26	234	99	58	38	13	5	7	1	0	0	0	0	0	0
27	192	138	74	39	9	2	1	0	0	0	0	0	0	0
28	198	134	65	46	8	3	1	0	0	0	0	0	0	0

29	226	98	67	45	13	0	6	0	0	0	0	0	0	0	
30	192	135	83	30	12	2	1	0	0	0	0	0	0	0	
31	220	101	70	54	4	0	6	0	0	0	0	0	0	0	
32	189	142	74	39	8	2	1	0	0	0	0	0	0	0	
33	185	141	82	40	5	2	0	0	0	0	0	0	0	0	
34	183	145	84	32	9	2	0	0	0	0	0	0	0	0	
35	187	140	88	24	13	3	0	0	0	0	0	0	0	0	
36	160	173	90	26	6	0	0	0	0	0	0	0	0	0	
37	150	173	114	18	0	0	0	0	0	0	0	0	0	0	
38	161	173	86	30	5	0	0	0	0	0	0	0	0	0	
39	172	172	65	34	9	3	0	0	0	0	0	0	0	0	
40	175	164	79	19	14	4	0	0	0	0	0	0	0	0	
41	178	152	91	18	13	3	0	0	0	0	0	0	0	0	
42	153	184	86	30	1	1	0	0	0	0	0	0	0	0	
43	159	171	97	23	4	1	0	0	0	0	0	0	0	0	
44	157	172	99	23	4	0	0	0	0	0	0	0	0	0	
45	160	174	87	29	5	0	0	0	0	0	0	0	0	0	
46	151	185	89	28	2	0	0	0	0	0	0	0	0	0	
47	169	166	81	31	6	2	0	0	0	0	0	0	0	0	
48	152	189	78	34	2	0	0	0	0	0	0	0	0	0	
49	149	189	87	28	2	0	0	0	0	0	0	0	0	0	
50	139	201	93	20	2	0	0	0	0	0	0	0	0	0	
51	160	178	79	33	5	0	0	0	0	0	0	0	0	0	
52	162	176	76	37	4	0	0	0	0	0	0	0	0	0	
53	171	163	82	29	9	1	0	0	0	0	0	0	0	0	
54	175	157	83	31	6	3	0	0	0	0	0	0	0	0	
55	162	172	84	33	4	0	0	0	0	0	0	0	0	0	
56	151	195	72	32	5	0	0	0	0	0	0	0	0	0	
57	136	202	98	19	0	0	0	0	0	0	0	0	0	0	
58	156	180	90	22	6	1	0	0	0	0	0	0	0	0	
59	187	149	61	51	4	3	0	0	0	0	0	0	0	0	
60	150	186	89	29	1	0	0	0	0	0	0	0	0	0	
61	245	35	126	42	0	0	7	0	0	0	0	0	0	0	
62	163	164	97	27	4	0	0	0	0	0	0	0	0	0	
63	151	185	88	30	1	0	0	0	0	0	0	0	0	0	
64	178	143	97	30	7	0	0	0	0	0	0	0	0	0	
65	150	188	86	29	2	0	0	0	0	0	0	0	0	0	
66	146	189	95	24	1	0	0	0	0	0	0	0	0	0	
67	138	203	90	24	0	0	0	0	0	0	0	0	0	0	
68	146	190	94	23	2	0	0	0	0	0	0	0	0	0	
69	143	190	101	21	0	0	0	0	0	0	0	0	0	0	
70	159	181	75	36	4	0	0	0	0	0	0	0	0	0	
71	141	196	96	21	1	0	0	0	0	0	0	0	0	0	
72	137	206	88	23	1	0	0	0	0	0	0	0	0	0	

73	136	211	82	24	2	0	0	0	0	0	0	0	0	0
74	144	205	72	30	4	0	0	0	0	0	0	0	0	0
75	144	195	90	25	0	1	0	0	0	0	0	0	0	0
76	165	170	80	35	5	0	0	0	0	0	0	0	0	0
77	123	224	93	15	0	0	0	0	0	0	0	0	0	0
78	134	215	78	28	0	0	0	0	0	0	0	0	0	0
79	126	213	108	6	2	0	0	0	0	0	0	0	0	0
80		60	365	0	30	0	0	0	0	0	0	0	0	0

4. Comparison.

In order to test these invariants, we have generated STS using a hill-climbing algorithm described in [10]. This algorithm will construct STS extremely quickly, and we hope that STS produced in this way will be (at least, to a degree) random. The average times taken (per design) are presented in Table 2. (The algorithms were programmed in PASCAL/VS and run on the University of Manitoba AMDAHL 470/V8 computer.)

Table 2. (time in seconds)

v	construction	fragment vector	indegree list
15	.011	.0021	.0095
19	.021	.0044	.022
21	.028	.0060	.030
25	.044	.010	.051
27	.048	.013	.065
31	.068	.017	.090

We see that the indegree sequence requires about five times as long to calculate as the fragment vector. Nevertheless, both are extremely quick.

We have noted that both fragment vectors and indegree lists are complete invariants for $v \leq 15$. The next step is to test their effectiveness on STS of order 19. Of 36000 designs constructed, 23966 distinct fragment vectors were found, and 32292 distinct indegree lists were obtained. Of the last 1000 designs constructed, there were 553 new fragment vectors and 785 new indegree lists.

These two invariants seem to be quite uncorrelated. Of the 36000 designs, there were only two that had the same fragment vector and indegree list.

So, it appears that indegree lists are both slower and more effective than fragment vectors. Both invariants seem to be very successful in practice. They can also easily be combined and used as single invariant.

We are currently investigating STS of order 19, using both these invariants. We have noted that there are known to be 284407 non-isomorphic STS of order 19; we hope to improve this lower bound using the methods described in this paper.

References.

[1] M.J. Colbourn, C.J. Colbourn and W.L. Rosenbaum, *Trains: an invariant for Steiner triple systems,* Ars Combinatoria 13 (1982), 149-162.

[2] F.N. Cole, *The triad systems of thirteen elements,* Trans. Amer. Math. Soc. 14 (1913), 1-5.

[3] L.D. Cummings, *On the method of comparison for triple systems,* Trans. Amer. Math. Soc. 15 (1914), 311-327.

[4] P.B. Gibbons, *Computing techniques for the construction and analysis of block designs,* Ph.D. Thesis, University of Toronto, 1976.

[5] T.P. Kirkman, *On a problem in combinations,* Cambridge and Dublin Math. Journal 2 (1847), 191-204.

[6] R.A. Mathon, K.T. Phelps and A. Rosa, *Small Steiner triple systems and their properties,* Ars Combinatoria 15 (1983), to appear.

[7] G.L. Miller, *On the $n^{\log n}$ isomorphism technique,* Proc. of the Tenth Annual ACM Symposium on the Theory of Computing, (1978), 51-57.

[8] L.P. Petrenjuk and A.J. Petrenjuk, *An enumeration method for non-isomorphic combinatorial designs,* Annals of Discrete Math., 7 (1980), 265-276.

[9] D.R. Stinson, *Isomorphism testing of Steiner triple systems: canonical forms,* pre-print.

[10] D.R. Stinson, *Hill-climbing algorithms for the construction of combinatorial designs,* in "Algorithms in Combinatorial Design Theory", Annals of Discrete Math., to appear.

[11] H.S. White, *Triple systems as tranformations, and their paths among triads,* Trans- Amer. Math. Soc. 14 (1913), 6-13.

Department of Computer Science
University of Manitoba
Winnipeg, Manitoba
R3T 2N2 Canada