COMBINATORICA

Erdős-Gallai test in linear time

Antal Iványi*
Department of Computer Algebra,
Eötvös Loránd University, Hungary
email: tony@compalg.inf.elte.hu

Loránd Lucz*
Department of Computer Algebra,
Eötvös Loránd University, Hungary
email: lorand.lucz@caesar.elte.hu

Abstract. Havel in 1955 [10], Erdős and Gallai in 1960 [6], Hakimi in 1962 [9], Ruskey, Cohen, Eades and Scott in 1994 [21], Barnes and Savage in 1997 [2], Tripathi, Venugopalan and West in 2010 [29] proposed a method to decide, whether a sequence of nonnegative integers can be the degree sequence of a simple graph (such sequences are called graphical). The running time of their algorithms in worst case is $\Omega(n^2)$. In this paper we propose a new algorithm called LED (Linear Erdős-Gallai algorithm), whose worst running time is $O(n)$. Since in the case of a graphical sequence all elements of the investigated sequence are to be tested, in the case of RAM model of computations LEG is asymptotically optimal.

1 Introduction

In practice we often need the ranking of different objects. One of the popular methods is pairwise comparison of the objects and assignment points to the objects, summing the assigned points and ranking of the objects according to the number of the gathered points. E. G. Landau referenced biological [16], Hakimi chemical, [9], Kim et al. [15], Newman and Barabási [20] network, Bozóki, Fülöp, and Rónyai economical [3] applications, Liljeros et al. applications connected with human connections [17], Iványi et al. connected with sport [11, 12, 14].

A számos lehetséges terminológia közül Erdős Pál és Gallai Tibor [6] cikkének szóhasználatát és jelöléseit követjük. A pontozás módjától függően sokféle feladat és eredmény van. Ebben a cikkben gráfon olyan véges, irányítás nélküli gráfot értünk, amelyben nem fordulnak elő hurokélek és bármely két csúcsot

Mathematics Subject Classification 2010: 05C85, 68R10

*The European Union and the European Social Fund have provided financial support to the project under the grant agreement no. TÁMOP 4.2.1/B-09/1/KMR-2010-0003.

legfeljebb egy él köt össze – azaz *egyszerű gráfokról* lesz szó. Ez a gráf annak a pontozásnak felel meg, amikor az összehasonlításoknak két lehetséges eredménye van: vagy mindkét objektum egy, vagy pedig mindkettő nulla pontot kap. Az összehasonlítások eredményét ábrázoló gráfban két csúcs pontosan akkor van összekötve, amikor a nekik megfelelő objektumok összehasonlítása során mindkét objektum egy pontot kapott.

A sok népszerű feladat közül azt vizsgáljuk, hogyan dönthető el gyorsan, hogy nemnegatív egészek $n \geq 1$ hosszúságú $b = (b_1, \ldots, b_n)$ sorozatához létezik-e olyan egyszerű gráf, amelynek b a foksorozata.

A hasonló feladatokkal kapcsolatban megjegyezzük, hogy mind az irányítatlan, mind pedig az irányított gráfokkal kapcsolatban az utóbbi néhány évben is számos cikk és könyvfejezet jelent meg (például [**?**, **?**, 29], illetve [**?**, **?**, 7, 11, 12]).

Legyen $n \geq m \geq 11$ és $b \geq$. Egész számok $b = (b_1, \ldots, b_n m$ sorozatát (m, n)-*korlátosnak* (röviden: korlátosnak) nevezzük, ha $0 \leq b_i \leq n-1$ minden $1 \leq i \leq m$ indexre. A $b = (b_1, \ldots, b_m)$ n-korlátos sorozatot (m, n)-*szabályosnak* (röviden: szabályosnak) mondjuk, ha $n - 1 \geq b_1 \geq \cdots \geq b_m \geq 0$. Egy (m, n)-szabályos sorozatot n-*megvalósíthatónak* (röviden: *jónak*) nevezünk, ha létezik olyan G egyszerű gráf, hogy G foksorozata $(b_1, b_2, \ldots, b_m, 0^{n-m})$, ahol 0^{n-m} $n-m$ darab nulla elemet jelent. Ha egy korlátos vagy szabályos sorozat esetén $m = n$, akkor az n paramétert csak egyszer írjuk ki.

A továbbiakban főleg szabályos sorozatokkal foglalkozunk. A definíciókban az alsó és felső korlátok azért szerepelnek, hogy ellenőrző algoritmusainkat megkíméljük a nyilvánvalóan nem jó sorozatok ellenőrzésétől, ezért ezek a megszorítások nem jelentik az általánosság megszorítását.

Cikkünk fő célkitűzése, hogy minél kisebb várható futási idejű algoritmust találjunk annak eldöntésére, hogy előírt b szabályos sorozathoz található-e olyan G egyszerű gráf, melynek foksorozata b. A probléma önmagában is érdekes, azonban számunkra külön motivációt jelent, hogy a feladat részfeladatként jelentkezik, amikor minél kisebb várható futási idejű algoritmust keresünk annak eldöntésére, hogy előírt sorozat lehet-e egy labdarúgó bajnokság pontsorozata [7, **?**, 13, **?**, **?**, **?**, 13].

Melléktermékként bővítettük az EIS [23] adatbázist: egyrészt a benne lévő ismert sorozatok új elemeivel, másrészt konkrét alkalmazásokkal kapcsolatos új sorozatok eleminek megadásával.

A cikkben egyrészt a vizsgált feladatot pontosan megoldó klasszikus és általunk javasolt új algoritmusokat, másrészt a feladat közelítő megoldását végző, új közelítő algoritmusokat vizsgálunk.

A cikk felépítése a következő. A bevezető első rész után a témakör klasszikus algoritmusait foglaljuk össze. A harmadik rész az új tesztelő algoritmusokat ismerteti, az negyedik részben a közelítő algoritmusok hibáját vizsgáljuk, míg az ötödikben a pontos algoritmusok futási idejét elemezzük.

A cikk eredményeinek egy része a magyar nyelvű [?] cikkből származik.

2 Classical precise theorems

Given sequence $b = (b_1, \ldots, b_n)$ of nonnegative integers the first i elements of the sequence we call *the head* of the sequence belonging to the index i, while the last $n-i$ elements of the sequence we call *the tail* of the sequence belonging to the index i.

2.1 Havel-Hakimi theorem

A feladat megoldására az első módszert Vaclav Havel cseh matematikus javasolta 1955-ben [10, 18].

Theorem 1 (Havel [10]) *If $n \geq 3$, then the n-regular sequence $b = (b_1, \ldots, b_n)$ is n-good iff the sequence $b' = (b_2 - 1, b_3 - 1, \ldots, b_{b_1} - 1, b_{b_1+1} - 1, b_{b_1+2}, \ldots, b_n)$ is $(n-1)$-good.*

Proof. See [10]. □

Ha ezen tétel alapján írunk egy rekurzív algoritmust, akkor annak futási ideje a RAM számítási modell [5] szerint legrosszabb esetben $\Theta(n^2)$. Érdemes megjegyezni, hogy a tétel bizonyítása konstruktív, és a bizonyításon alapuló algoritmus négyzetes idő alatt nem csak ellenőriz, hanem egy megfelelő gráfot is előállít (feltéve persze, hogy létezik megfelelő egyszerű gráf). Mivel például a teljes gráfnak $\binom{n}{2}$ éle van, a megvalósítást is biztosító algoritmus futási ideje nem lehet kisebb, mint négyzetes.

2.2 Erdős-Gallai theorem

In 1962 Louis Hakimi [9] published independently the same result, therefore the theoren is called today usually as *Havel-Hakimi theorem*, and the method of reconstruction is called *Havel-Hakimi algorithm*.

Időrendben a következő eredmény Erdős Pál és Gallai Tibor alábbi szükséges és elégséges feltétele [6] volt.

Theorem 2 (Erdős, Gallai, [6]) *Let $n \geq 3$. The n-regular sequence $b = (b_1 \ldots, b_n)$ is n-good iff*

$$\sum_{i=1}^{n} b_i \quad \textit{páros} \tag{1}$$

and

$$\sum_{i=1}^{j} b_i - j(j-1) \leq \sum_{k=j+1}^{n} \min(j, b_k) \quad (j = 1, \ldots, n-1). \tag{2}$$

Proof. See [**?**, 6, 22, 29]. □

Bár ez a tétel csak ellenőriz, (2) módszeres alkalmazása *minden* esetben $\Theta(n^2)$ időt igényel. A közelmúltban

Tripathi et al. [29] publikáltak a tételre konstruktív bizonyítást, amely megvalósítható bemenet esetén $O(n^3)$ idő alatt egy megoldást is előállít.

3 Linear time Erdős-Gallai test

A következő tétel még a EG-FAROKFELEZŐ algoritmusnál is jobban kihasználja, hogy a b bemeneti sorozat monoton. Legyen $h = \min\{k \mid b_k \leq i,\ i = 1, \ldots, n\}$.

Theorem 3 *Ha $n \geq 1$, az n-szabályos $b = (b_1 \ldots, b_n)$ sorozat akkor és csak akkor n-jó, ha*

$$H_n \quad \textit{páros} \tag{3}$$

és

$$H_i - i(i-1) \leq (H_n - H_h) + (i + 1 - h)(h-1)/2 \quad \textit{páros} \tag{4}$$

Proof. Mivel a b_i sorozat csökkenő, az i sorozat pedig növekvő, van metszéspontjuk, amely előtt az indexeke, és amely után a b sorozat elemeinek összege adja a megfelelő farokrész kapacitását. □

Corollary 4 *Nemnegatív egészek n-szabályos $b = b_1, \ldots, , b_n$ sorozatáról $O(n)$ idő alatt eldönthető, hogy létezik-e olyan G egyszerű gráf, amelynek b a foksorozata.*

Proof. A 3 tételben szereplő h "metszéspont" $O(n)$ (sőt, akár $O(\log n)$) idő alatt meghatárározható (a metszéspontot minden sorozatra csak egyszer kell meghatározni.) Ezek után a b sorozat minden elemére $O(1)$ idő alatt elvégezhető a (4) egyenlőtlenség ellenőrzése. □

Érdemes megjegyezni, hogy akár a bemenő sorozat rendezettségétől is eltekinthetünk. Mivel a sorozat elemei egész számok és mindegyik a $[0, n-1]$ intervallumba esik, szükség esetén $O(n)$ idő alatt rendezni tudjuk.

A következő program a . tétel alapján adott n−-re az összes n-páros sorozatot előállítja és teszteli, hogy jók-e. A teljes program futási ideje minden sorozatra $(O(n)$, így a teljes futási idő $O(n\epsilon(n))$.

Bemenet. n: number of vertices $(n \geq 1)$;
$b = (b_1, \ldots, b_n)$: n-regular sequence.

Kimenet. L: logical variable, whose value is True, if the input is graphical, and False otherwise.

Working variables. i and j: cycle variables;
$H = (H_1, \ldots, H_n)$: H_i is the sum of the first i elements of b;
h: the intersection point of the input sequence b.

```
EG-Linear(n, b, L)
01 H_1 = 0                              ▷ Line 01: initialization
02 for i = 2 to n                       ▷ Lines 02–03: computation of the elements of H
03     H_i = H_{i-1} + b_i
04     i = 1                            ▷ Lines 04–07: computation of h
05     while i < b_i
06         i = i + 1
07     h = i
08     for i = 1 to n                   ▷ Lines 08–18: test of b
09         while i ≤ h
10             if i < h
11                 if H_i − i(i − 1) > (h − i)(h − i − 1) + H_n − H_h
12                 L = False
13                 return L
14             if i ≥ h
15                 if H_i − i(i − 1) > H_n − H_i
16                 L = False
17                 return L
18                 go to
19     L = True                         ▷ Lines 19–20: program ends with True
20     return L
```

Acknowledgements. The authors thank Professor András Frank (Eötvös Loránd University) for the continuous consultations.

n	$\epsilon(n)$	$\tau(n)$, mp	Op(n)	$\tau(n)/\epsilon(n)$, mp	$Op(n)/\epsilon(n)$
1	???	???	???	???	???
2	???	???	???	???	???
3	???	???	???	???	???
4	???	???	???	???	???
5	???	???	???	???	???
6	???	???	???	???	???
7	???	???	???	???	???
8	???	???	???	???	???
9	???	???	???	???	???
10	???	???	???	???	???
11	???	???	???	???	???
12	???	???	???	???	???
13	???	???	???	???	???
14	???	???	???	???	???
15	???	???	???	???	???
16	???	???	???	???	???
17	???	???	???	???	???
18	???	???	???	???	???
19	???	???	???	???	???
20	???	???	???	???	???
21	???	???	???	???	???
22	???	???	???	???	???
33	???	???	???	???	???
24	???	???	???	???	???
25	???	???	???	???	???

Figure 1: Total running time of EG-Linear for all regular sequences in number of operations, and running time per processed elements for all regular sequences in sec and in number of operations.

References

[1] T. M. Barnes, C. D. Savage, A recurrence for counting graphical partitions, *Electron. J. Combin.* **2,** (1995), Research Paper 11, 10 pages (electronic). ⇒

[2] T. M. Barnes, C. D. Savage, Efficient generation of graphical partitions. *Discrete Appl. Math.* **78,** (1–3) (1997) 17–26. ⇒1

[3] Bozóki S.: J. Fülöp, L. Rónyai: On optimal completion of incomplete pairwise comparison matrices. *Math. Comput. Modelling* **52,** (2010) 318–333. ⇒1

[4] Burns, J. M.: *The number of degree sequences.* PhD Dissertation, MIT, 2007. ⇒

[5] Cormen, T. H., Leiserson, Ch. E., Rivest, R. L., Stein, C.: *Introduction to Algorithms.* Third edition, The MIT Press/McGraw Hill, Cambridge/New York, 2009. ⇒3

[6] Erdős, P., Gallai, T.: *Gráfok előírt fokú pontokkal.* Mat. Lapok **11,** (1960) 264–274. ⇒1, 3, 4

[7] Frank, A.: *Connections in Combinatorial Optimization.* Oxford University Press, Oxford, 2011. ⇒2

[8] Frank, D. A., Savage, C. D., Sellers, J. A.: *On the number of graphical forest partitions. Ars Combin.* **65** (2002) 33–37. ⇒

[9] Hakimi, S. L.: *On the realizability of a set of integers as degrees of the vertices of a simple graph.* J. SIAM Appl. Math. **10,** (1962) 496–506. ⇒ 1, 3

[10] Havel, V.: A remark on the existence of finite graphs (cseh); *Časopis Pěst. Mat.* **80,** (1955), 477–480. ⇒1, 3

[11] Iványi, A.: *Reconstruction of complete interval tournaments.* Acta Univ. Sapientiae, Informatica, **1(1),** (2009) 71-88. ⇒1, 2

[12] Iványi, A.: *Reconstruction of complete interval tournaments. II.* Acta Univ. Sapientiae, Mathematica, **2(1),** (2010) 47–71. ⇒1, 2

[13] Iványi, A.: *Deciding the validity of the score sequence of a soccer tournament.* In: Open problems of the Egerváry Research Group, Budapest, 2011. ⇒2

[14] Iványi, A., Lucz, L., Sótér, P.: *On the Erdős-Gallai and Havel-Hakimi algorithms.* Acta Univ. Sapientiae, Informatica (benyújtva). ⇒ 1

[15] Kim, H., Toroczkai, Z., Miklós, I., Erdős, P. L., Székely, L. A.: Degree-based graph construction. J. Physics: Math. Theor. A **42(39),** (2009) 392–401. ⇒1

[16] Landau, H. G.: On dominance relations and the structure of animal societies. III. The condition for a score sequence. *Bull. Math. Biophys.* **15,** (1953) 143–148. ⇒1

[17] LILJEROS, F., EDLING, C. R., AMARAL, L., STANLEY, H., ÁBERG, Y.: *The web of human sexual contacts.* Nature **411,** (2001) 907–908. ⇒ 1

[18] LOVÁSZ, L.: *Combinatorial Problems and Exercises* (corrected version of the second edition). AMS Chelsea Publishing, Boston, 2007. Magyarul: *Kombinatorikai problémák és feladatok.* Typotex, Budapest, 1999. ⇒ 3

[19] MOON, J. W.: *Topics on Tournaments.* Holt, Rinehart, and Winston, New York, 1968. ⇒

[20] NEWMAN, M. E. J., BARABÁSI, A. L.: *The Structure and Dynamics of Networks.* Princeton University Press, Princeton, NJ. 2006. ⇒ 1

[21] RUSKEY, F., COHEN, R., EADES, P., SCOTT, A.: *Alley CAT's in search of good homes.* Congressus Numerantium, **102** (1994) 97–110. ⇒ 1

[22] SIERKSMA, G., HOOGEVEEN, H.: *Seven criteria for integer sequences being graphic.* J. Graph Theory **15(2),** (1991) 223–231. ⇒ 4

[23] SLOANE N. J. A. (szerkesztő): *Encyclopedia of Integer Sequences.* 2011. ⇒ 2

[24] SLOANE N. J. A.: *The number of ways to put $n+1$ indistinguishable balls into $n+1$ distinguishable boxes.* In: The On-line Encyclopedia of the Integer Sequences. 2011. ⇒

[25] SLOANE N. J. A.: *The number of degree-vectors for simple graphs.* In: The On-line Encyclopedia of the Integer Sequences. 2011. ⇒

[26] SLOANE N. J. A.: *The number Number of bracelets with n red, 1 pink and $n-1$ blue beads.* In: The On-line Encyclopedia of the Integer Sequences. 2011. ⇒

[27] STANLEY, R.: *A zonotope associated with graphical degree sequence.* In: Applied geometry and discrete mathematics, Festschr. 65th Birthday Victor Klee. DIMACS Series in Discrete Mathematics and Theoretical Computer Science. **4,** (1991) 555-570. ⇒

[28] TRIPATHI, A., VIJAY, S.: *A note on a theorem of Erdős & Gallai.* Discrete Math. **265(1–3),** (2003) 417–420. ⇒

[29] TRIPATHI, A., VENUGOPALANB, S., WEST, D. B.: *A short constructive proof of the Erdős-Gallai characterization of graphic lists.* Discrete Math. **310(4),** (2010) 833–834. ⇒ 1, 2, 4