

ГЕОКОДИРАЩА СИСТЕМА P4

Павел Петров¹, Стефка Петрова²

¹ Икономически университет – Варна/Катедра „Информатика“, Варна, България, petrov@ue-varna.bg

² Колеж по туризъм/Катедра Туризъм, Варна, България, s.petrova@ue-varna.bg

РЕЗЮМЕ

В доклада се предлага нов начин за геокодиране, предназначен да преодолее някои от недостатъците, налични в съществуващите геокодиращи системи - системата P4. При тази система се използват правоъгълни клетки, на които се присвоява буквено-цифров код с дължина, зависеща от нивото на подразделяне (т.е. размерите) на клетката. Всички клетки от едно и също ниво на подразделяне имат еднакви площи. Специално внимание е отделено на полярните райони, които се обхващат от две специални клетки. Системата предлага различни нива на точност и може да се използва като заместител на традиционни географски данни, като например улични адреси или пощенски кодове.

КЛЮЧОВИ ДУМИ: геокодиране, P4, OLC/PC, карти, адреси

GEOCODING SYSTEM P4

Pavel Petrov¹, Stefka Petrova²

¹ University of Economics - Varna/Department of Informatics, Varna, Bulgaria, petrov@ue-varna.bg

² College of Tourism/Department of Tourism, Varna, Bulgaria, s.petrova@ue-varna.bg

ABSTRACT

The P4 system, proposed in the study, is a new method of geocoding that aims to address some of the disadvantages of earlier systems. This system employs rectangular cells that are assigned an alphanumeric number whose length is determined by the cell's level of subdivision (i.e., cell dimensions). All cells within the same subdivision level have the same area. Special attention is paid to the polar regions, which are covered by two special cells. The system offers different levels of accuracy and can be used as a substitute for traditional geographic data such as street addresses or postal codes.

KEYWORDS: geocoding, P4, OLC/PC, maps, addresses

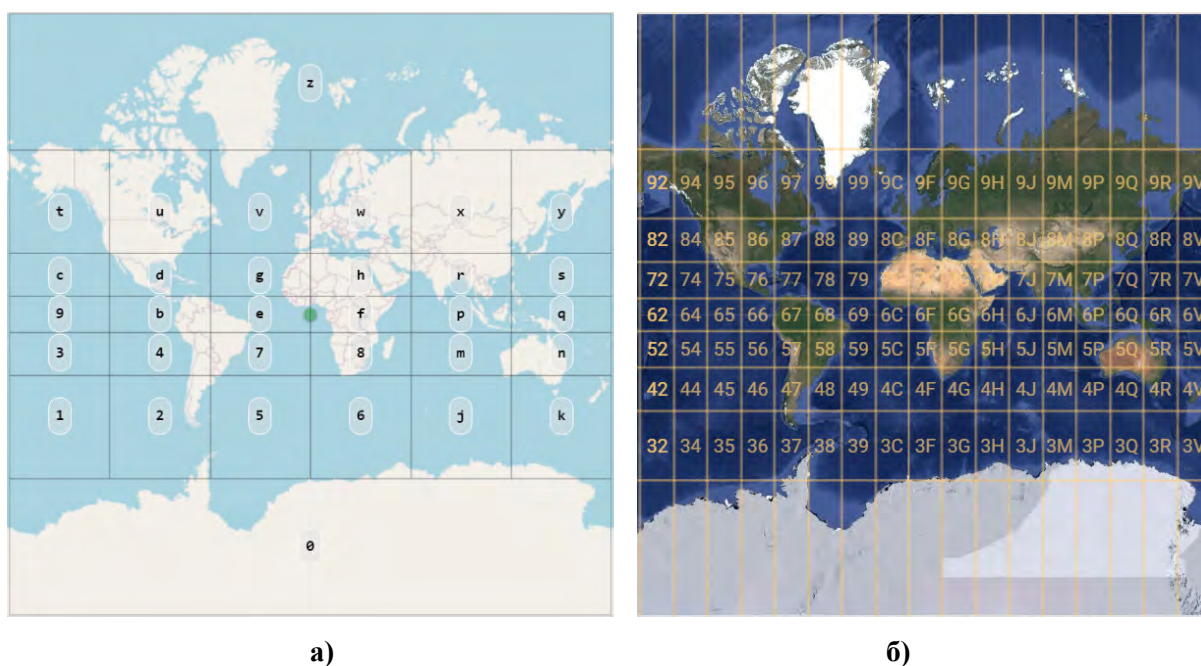
ВЪВЕДЕНИЕ

Системите за геокодиране имат важна роля при обработка на геопространствени данни, тъй като предоставят начин за кодиране на географски местоположения във формати, удобни, както за хората, така и за бърза компютърна обработка. Традиционната двойка за задаване на координати – географска ширина и дължина, имат универсално приложение, но не са подходящи за масова употреба, когато става въпрос за бързо и лесно определяне на приблизителни местоположения (Cliquet, 2021). В последните години бяха разработени различни системи за геокодиране като MGRS (Office of Geomatics, 2022), Quadkey, Geohash (Niemeyer, 2024), S2 (Google 2024b), H3 (Brodsky, 2018), OLC/PC (Google, 2024a) и мн. други с цел да се улесни кодирането на местоположения. Много от тези системи, според нас, още при създаването си имат редица недостатъци, което е и причината постоянно да възникват нови системи, които се опитват да решат едно или друго ограничение.

Системата Р4, която тук предлагаме, прилага нов подход за геокодиране, предназначен да преодолее някои от недостатъците, налични в съществуващите геокодиращи системи. В системата Р4 се използват правоъгълни клетки, на които се присвоява буквено-цифров код с дължина, зависеща от нивото на подразделяне (т.е. размерите) на клетката. Всички клетки от едно и също ниво имат еднакви площи. Специално внимание е отделено на полярните райони, които се обхващат от две специални клетки. Системата предлага различни нива на точност и може да се използва като заместител на традиционни географски данни, като улични адреси или пощенски кодове.

1. МЕТОДОЛОГИЯ И ДИЗАЙН НА СИСТЕМАТА Р4 НА ПЪРВО НИВО НА РАЗДЕЛЯНЕ НА ЗЕМНАТА ПОВЪРХНОСТ НА КЛЕТКИ

В системата за геокодиране Р4 се използват положителните характеристики на редица съществуващи геокодиращи системи, като се цели да се избегнат техни недостатъци. Системата Р4 (HashGeo, 2024) използва 32 символа за кодирането, подобно на системата Geohash-EAS (Petrov et al., 2018), както и същата схема на подразделяне на клетките, но от второ ниво нататък. Първоначалното разделяне на земната повърхност се извършва по различен начин на 7 реда – 2 клетки за полярните области (фиг. 1а), а между тях решетка от 5 реда и 6 колони ($2 + 5 \times 6 = 32$ клетки). За сравнение, системата OLC/PC на Google представлява решетка от 9 реда и 18 колони, като не се отделя специално внимание на полярните области (фиг. 1б).



Фигура 1. Първоначално разделяне на земната сферична повърхност:

а) 32 клетки с еднакви площи в геокодиращата система Р4;

б) 162 клетки с различни площи, но с еднакви размери в градуси – 20 на 20 градуса, в геокодиращата система OLC/PC.

Източник: а) Собствена разработка, <https://hashgeo.free.bg/explore/P4/>; б) Google, <https://plus.codes/map>

В системата OLC/PC последващите четири рекурсивни деления на всяка клетка са 20 на 20 клетки, т.е. по отделните нива клетките са с размери $20^{\circ} \times 20^{\circ}$, $1^{\circ} \times 1^{\circ}$, $0,05^{\circ} \times 0,05^{\circ}$, $0,0025^{\circ} \times 0,0025^{\circ}$ и $0,000125^{\circ} \times 0,000125^{\circ}$, при което площите на пространствата от едно и също ниво, но намиращи се на различни географски ширини, не са еднакви. Съответно, широчините (хоризонталния размер) на клетките от едно и също ниво, мерени в метри или километри също варират. Близко около екватора размерите на клетките от първите 5 нива са съответно приблизително 2224×2224 км, 111×111 км, $5,6 \times 5,6$ км, 278×278 м и $13,9 \times 13,9$ м. – те имат приблизителните размери на квадрати (приемаме, че на екватора $1^{\circ} = 111,195$ км. и Земята има сфероидна форма). Следващите деления са 5×4 клетки (размери край екватора около: $3,5 \times 2,8$ м, 87×56 см и т.н.) и формата се променя.

Можем да използваме следната формула за дължина на паралела на дадена географска ширина, за да определим измененията в размерите на клетките, разположени на различни географски ширини:

$$L = 2\pi R^2 \cos(\varphi) \quad (1),$$

където:

R – радиус на Земята – приемаме 6371,0088 км, съгласно WGS за аритметичен среден земен радиус, използван в GPS (Hofmann-Wellenhof et al., 2012)

φ – географска ширина

Използвайки формула (1) получаваме, че широчините на клетките се променят, както е показано в табл. 1:

Таблица 1.

Широчини на клетки от различни нива на различни географски ширини в OLC/PC

Ниво	10°S/N	30°S/N	50°S/N	70°S/N
1	2190 км	1926 км	1430 км	761 км
2	110 км	96 км	72 км	38 км
3	5,5 км	4,8 км	3,6 км	1,9 км
4	274 м	241 м	179 м	95 м

Както се вижда, широчините на клетките, разположени на географска ширина 70°S/N са около три пъти по-малки спрямо тези, разположени близо до екватора. Това рефлектира и върху площите на клетките – те също са около три пъти по-малки. За разлика от OLC/PC, в системата P4 се избягва промяната на площите, докато промяната на широчините не може да се избегне, поради сферичната форма. Запазването на еднаквостта на площите се постига, като височината на клетките на всеки ред се променя по такъв начин, че да компенсира намаляването на широчината на клетките (табл. 2).

Таблица 2.

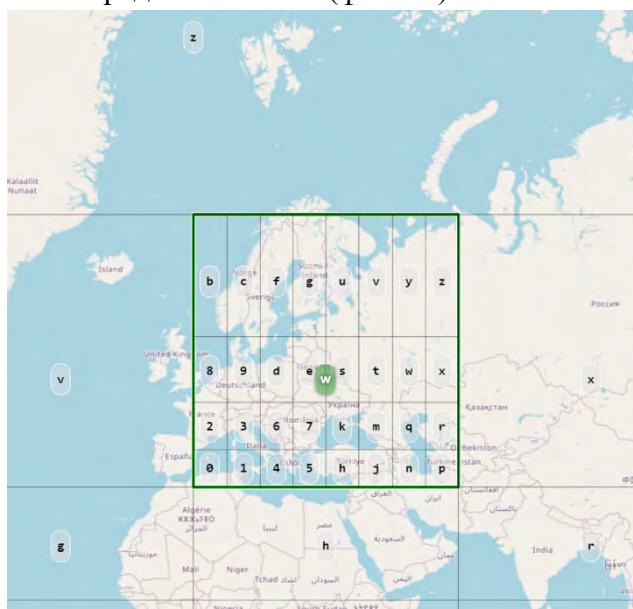
Хоризонтални граници между клетките, разположени на различните редове и техните височини в градуси

№ на ред	Клетки	Географска ширина на южна граница	Географска ширина на северна граница	Височина в градуси
1	0	90°S	69,6°S	20,4°
2	1, 2, 5, 6, j, k	69,6°S	34,2°S	35,4°
3	3, 4, 7, 8, m, n	34,2°S	10,8°S	23,4°
4	9, b, e, f, p, q	10,8°S	10,8°N	21,6°
5	c, d, g, h, r, s	10,8°N	34,2°N	23,4°
6	t, u, v, w, x, y	34,2°N	69,6°N	35,4°
7	z	69,6°N	90°N	20,4°
ОБЩО:				180°

Докато при OLC/PC тази височина е фиксирана на 20° при всички редове на първо ниво, в системата P4 височината на клетките варира от 21,6° до 35,4°, като специалните полярни клетки са с височина 20,4°.

2. ДИЗАЙН НА СИСТЕМАТА P4 НА ВТОРО И СЛЕДВАЩИ НИВА

В системата P4 на второ ниво клетките се разделят на 4 реда и 8 колони (фиг. 2а), на трето ниво - на 8 реда и 4 колони и т.н. В сравнение със системите Geohash-EAS и Geohash, схемата е обърната - при тях на четните нива клетките се разделят на 8 реда и 4 колони, а на нечетните нива - на 4 реда и 8 колони (фиг. 2б).



а)



б)

Фигура 2. Подразделяне на второ ниво в различни системи:

а) клетка „w” на 4 реда и 8 колони в геокодиращата система P4;

б) клетка „u” на 8 реда и 4 колони в геокодиращата система Geohash-EAS.

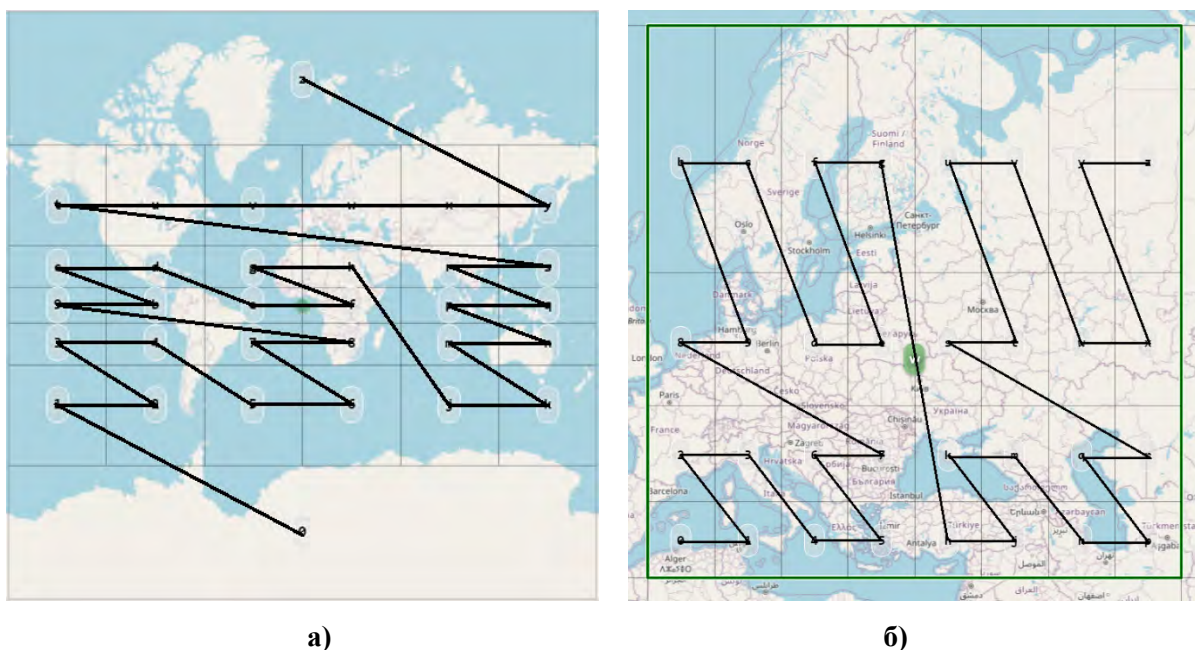
Източник: Собствена разработка, <https://hashgeo.free.bg/explore/P4/>

При системата P4 се запазва Z-кривата на подреждане на клетките, като на първо ниво системата е значително променена, тъй като броят клетки по редове и колони е друг, а на следващите нива е същата, както при Geohash-EAS, съответно при класическия Geohash (фиг. 3).

Както е известно, в системите за геокодирание Z-кривата се използва, за да се преобразуват пространствени данни от двумерно пространство в едномерно представяне, като се запази пространствената им близост. Характерно при този начин на разделяне е, че номерирането на клетките следва т.нар. Z-крива на обхождане. При геокодирането това значи, че клетките, които са близко една до друга в двумерното пространство, също да са близко една до друга в едномерно пространство. Разбира се, за отделни случаи на определени двойки клетки това не е изпълнено (например прехода "c-t"), но като цяло преходите между съседните клетки са минимизирани.

По наше мнение, тази особеност не се използва активно в практическите реализации и в този смисъл преминаването от една клетка в друга, независимо дали се извършва по Z-крива на обхождане (известна още и като крива на Мортон и крива на Лъобег), по крива на Хилберт, по крива на Мур, по крива на Шерпински или по някоя друга от кривите на Пеано, има по-скоро теоретично значение.

Според нас, в редица геокодиращи системи не се отчита фактът, че при първоначалното деление, клетките, разположени най-ляво и най-дясно, всъщност са съседни, тъй като граничат с 180-градусовия меридиан. Това налага, поне при първоначалното деление, прилагането на различен подход, което считаме, че трябва да е обект на бъдещи изследвания по темата.



Фигура 3. Обхождаща Z-крива в геокодиращата система P4:

а) на първо ниво на подразделяне;

б) на второ ниво на подразделяне – клетка „w“.

Източник: Собствена разработка

ЗАКЛЮЧЕНИЕ

Геокодиращата система Р4 предлага нов подход за решаване на някои от недостатъците на съществуващите системи за геокодиране. Тя използва правоъгълни клетки с равни площи, като особено внимание е отделено на полярните райони. На първо ниво системата Р4 разделя земната повърхност на 32 клетки с различни височини, което осигурява равноплътност на всички клетки, за разлика от други системи като OLC/PC, при които площите на клетките варират според географската ширина. На следващи нива системата Р4 следва рекурсивно подделение, като запазва Z-кривата за минимизиране на преходите между съседни клетки. Системата цели да предложи по-ефективно представяне на пространствени данни. Системата може да послужи за основа при бъдещи изследвания в областта на оптимизацията при първоначалното деление, както и на специалните случаи като полярните области и граничните клетки.

REFERENCES / ИЗПОЛЗВАНА ЛИТЕРАТУРА

1. Brodsky, I. (2018) H3: Uber's Hexagonal Hierarchical Spatial Index [Online] Available from: <https://www.uber.com/blog/h3/> [Accessed 10/10/2024].
2. Cliquet, G. (2021) From Geomarketing to Spatial Marketing. Spatial Economics, Volume II: Applications, pp.277-305.
3. Google (2024a) Open Location Code Documentation. [Online] Available from: <https://plus.codes> [Accessed 10/10/2024].
4. Google (2024b) S2 Geometry Library [Online] Available from: <https://s2geometry.io> [Accessed 10/10/2024].
5. HashGeo (2024) Exploration of the P4 Geocoding System. [Online] Available from: <https://hashgeo.free.bg/explore/P4> [Accessed 10/10/2024].
6. Hofmann-Wellenhof, B., Lichtenegger, H., & Collins, J. (2012). Global positioning system: theory and practice. Springer Science & Business Media.
7. Niemeyer, G. (2024) Geohash [Online] Available from: <https://geohash.org> [Accessed 10/10/2024].
8. Office of Geomatics (2022) Military Grid Reference Systems [Online] Available from: <https://earth-info.nga.mil/index.php?dir=coordsys&action=coordsys> [Accessed 10/10/2024].
9. Petrov, P., Dimitrov, P. & Petrova, S. (2018). Geohash-EAS – a Modified Geohash Geocoding System with Equal-Area Spaces. SGEM 2018 : 18 International Multidisciplinary Scientific Geoconference : Conference Proceedings, Sofia : STEF92 Technology Ltd., Vol.18 Informatics, Geoinformatics a. Remote Sensing, Iss. 2.2, pp.187-194. DOI 10.5593/sgem2018/2.2/S08.024