

ISSN 1728-7901



Абай атындағы
Қазақ ұлттық педагогикалық университеті
Казахский национальный педагогический
университет имени Абая

ХАБАРШЫ ВЕСТНИК



Серия «Физико-математические науки» • «Физика-математика ғылымдары» сериясы

Алматы

№ 1 (49)
2015

Казахский национальный
педагогический университет
имени Абая
ВЕСТНИК
серия "Физико-математические
науки" № 1 (49)

Главный редактор
Академик НАН РК Г.У. Уалиев

Редакционная коллегия:

зам.главного редактора:

д.п.н. Е.Ы. Бидайбеков,

к.ф.-м.н. М.Ж. Бекпатшаев

ответ.секретарь

к.п.н. Г.А. Абдулкаримова

члены:

Dr.-ing. Holm Altenbach (Germany),

Dr. S.A.Hasan (Pakistan),

Dr. Yasuhide Fukumoto (Japan),

Phd.d Shuo-Hung Chang, (Taiwan),

д.п.н. А.Е. Абылкасымова,

д.ф.-м.н. М.А. Бектемесов,

д.ф.-м.н. А.С.Бердышев,

д.п.н. В.В. Гриншкун (Россия),

к.ф.-м.н. Ф.Р. Гусманова,

д.т.н. А.Д.Джураев (Узбекистан),

д.ф.-м.н. С.И. Кабанихин (Россия),

д.ф.-м.н. Б.А. Кожамкулов,

д.ф.-м.н. В.Н. Косов,

д.ф.-м.н. К.К. Коксалов,

д.т.н. М.К. Кулбеков,

д.п.н. М.П. Лапчик (Россия),

д.ф.-м.н. Қ.М. Мукашев,

д.ф.-м.н. С.Т. Мухамбетжанов,

д.т.н. Г.Я. Пановко (Россия),

д.п.н. Б.Д. Сыдыков,

д.ф.-м.н. Н.Ж. Такибаев,

д.ф.-м.н. К.Б. Тлебаев,

д.т.н. А.К. Тулешов,

д.ф.-м.н. З.Г. Уалиев,

д.ф.-м.н. Л.М. Чечин,

к.ф.-м.н. Е.Б. Шалбаев,

к.т.н. Ш.И. Хамраев

© Казахский национальный
педагогический университет
им. Абая, 2015

Зарегистрирован в Министерстве
информации Республики Казахстан,
№ 4824 - Ж - 15.03.2004
(периодичность—4 номера в год)
Выходит с 2000 года

Редакторы: Ф.Р. Гусманова,
Г.А. Абдулкаримова

Компьютерная верстка:
Г.А. Абдулкаримова
Ф.Р. Гусманова

Подписано в печать 27.03.2015 г.
Формат 60х84 1/8.
Об 13,75 уч.-изд.л.
Тираж 300 экз.

050010, г.Алматы, пр.Достык, 13,
КазНПУ им.Абая
Отпечатано в типографии
"TOO Palitra Press"
г.Алматы, ул.Хамиди 4а

П.О. Бердиева, Л.М. Чечин Об особенностях преподавания квантовой космологии в курсе «Современная астрономия»	111
К. Бисембаев, П. Кудайбергенкызы Колебания виброзащитных устройств на опорах качения, ограниченных поверхностями вращения высокого порядка при наличии трения качения на упруговязких грунтах	115
К. Бисембаев, Т.Б. Дикамбай Колебания в кулачковом механизме при законе изменения ускорения толкателя по параболе высокого порядка	122
М. Құлбекұлы, С. Жолдасбекова, Д.М. Кулбеков Күрделі кабаттаса жүретін тасымалдау үдерістерінің динамикасын зерттеп оқып-үйренудің кейбір мәселелері	130
Қ.Б. Тлебаев, Ә.Қ. Шоқанов, С.А. Шомшекова Атомдық күштік микроскоп арқылы политеграфторэтиленнің беткі қабатына зерттеулер жүргізу	135
К.Б. Тлебаев, А.О. Нусипова Исследование поверхности полиэтилентерефталата	139

ИНФОРМАТИКА. ИНФОРМАТИКАНЫ ОҚИТУ ӘДІСТЕМЕСІ. БІЛІМ БЕРУДІ АҚПАРАТТАНДЫРУ ИНФОРМАТИКА, МЕТОДИКА ПРЕПОДАВАНИЯ ИНФОРМАТИКИ.ИНФОРМАТИЗАЦИЯ ОБРАЗОВАНИЯ

А.Ж. Асаинова, Г.С. Джарасова Предметная область профессиональной подготовки магистра информатики	144
Е.Ы. Бидайбеков, К.Т. Алдияров Информационно- коммуникационные технологии как средство повышения профессионального мастерства преподавателей специальных дисциплин	152
Е.Ы. Бидайбеков, К.Т. Алдияров Использование электронных ресурсов по общетехническим дисциплинам в политехническом колледже	157
F.R. Gusmanova, A.S. Dossym, M.G. Sakipbekova Modelling of calculating processes of grid-system by natural algorithms of computation	161
F.R. Gusmanova, N.B. Ospanova, A. Altybai Continuation of nonlinear traffic time, to solve the task using macroscopic model of transportation	164
Г.С. Джарасова, А.Ж. Асаинова Компетентностная модель подготовки магистра информатики	169
М.Н. Калимолдаев, Ш.А. Малбасова Білім беру порталдарын жасау мен дамыту саласындағы педагогикалық зерттеулер	176
М.Н. Калимолдаев, М.С. Әшім Электрондық құжаттарды басқаратын ақпараттық жүйелерді талдау	182
С.М. Кенесбаев, А.К. Оралбекова Инклюзивті білім беруде ақпараттық және коммуникациялық технологияларды қолдану мәселелері	188
Н.С. Кольева, С.С. Жекеева Виртуалды білім беру кеңістігінде оқушылардың ақпараттық құзыреттілігін дамытудың мәселелері	192
Нугманова, Т.Ж. Мазаков, Г.С. Байрбекова О тенденции и развитии современных биометрических технологий	198
А.С. Омарбекова, А.Б. Закирова, А.Ф. Турсумбаева, К.Ж. Садваксова Методология разработки портала электронных учебных изданий	203
Ж.Е. Темирбекова, М.Е. Мансурова К MEANS алгоритмін JAVA MPJ EXPRESS және HADOOP технологиясында салыстыру	208
Ш.Т. Шекербекова, И. Бодаева Информатиканы медициналық жоғары оқу орындарында оқытудың ерекшеліктері	214

**ИНФОРМАТИКА. ИНФОРМАТИКАНЫ ОҚЫТУ ӘДІСТЕМЕСІ.
БІЛІМ БЕРУДІ АҚПАРАТТАНДЫРУ
ИНФОРМАТИКА, МЕТОДИКА ПРЕПОДАВАНИЯ ИНФОРМАТИКИ
ИНФОРМАТИЗАЦИЯ ОБРАЗОВАНИЯ**

1. Sharipbay A.A., Seifullina A., Omarbekova A.S., Zakirova A.B. Creating Intelligent Electronic Textbooks based on Knowledge Management/ The World Congress on Engineering 2014 (WCE 2014). Volume 1, 2014, P.224-227.
2. Sharipbay A.A., Omarbekova A.S. Generator of electronic educational publications. Seattle-2013: 4th International Academic Research Conference on Business, Education, Nature and Technology. 2013/Seattle, USA.- p.350-354
3. Омарбекова А.С. Турсумбаева А.Ф., Портал коммерциализации генератора электронных учебных изданий. Информатизация общества: Труды IV международной научно-практической конференции.-Астана, 2014.-С.278-281.

Аңдатпа. Бұл мақалада электронды оқу құралдарының қажеттілігі туралы сұрақ қарастырылады (ЭОҚ). Сонымен қатар порталды жасаудың әдіснамасы келтірілген. Өңделген порталдың жүйеасты оқытушыда, білім алушыда, администраторда болады, ол оқытушыға ЭОҚ генераторын, дайын ЭОҚ серверге жүктеуге және ЭОҚ сатуда пайызын алуға мүмкіндік береді.

Түйін сөздер: портал, электронды оқу құралы, электронды оқу құралының генераторы.

Abstract. This article discusses about the necessity for the portal of electronic educational editions (EEE). And also it shows the development of the methodology of this portal. The designed portal has a subsystem of the teacher, student, administrator, which allows teachers to download generator EEE, download ready EEE to the server and receiving a percentage of sales EEE.

Keywords: the portal, electronic educational editions, the generator of electronic educational editions.

ӨОЖ 004.75

Ж.Е. Темирбекова, М.Е. Мансурова

**K MEANS АЛГОРИТМІН JAVA MPJ EXPRESS ЖӘНЕ HADOOP
ТЕХНОЛОГИЯСЫНДА САЛЫСТЫРУ**

(Алматы қ., әл-Фараби атындағы Қазақ ұлттық университеті)

Аңдатпа. *k means* алгоритмі кластерлеу әдістерінің ішіндегі ең кең таралған *k means* Алгоритм қарапайымдылығы және орындалу жылдамдығы жөңары болғандықтан *k means* қолданысқа ие. Жұмыста гиперспектральды бейінені өңдеу, параллельді есептеу, үлестірілу есептеу туралы жұмыстар орындалды. Бағдарлама жиынында Java бағдарламаның ортасында MPJ Express кітапханасын қолдану арқылы және Hadoop технологиясында *k means* алгоритмі арқылы әр объектіні анық етіп бөліп кластерлеу қарастырылды.

Түйін сөздер: кластерлеу, параллельді есептеу тиімділігі, үлестірілу бағдарлама.

Өртүрлі спектралды ауқымда түсірілген спутник суреттері өте пайдалы мәліметтер тұрады және сандық түрде сақталады. Космостық суреттер орташа және үлкен масштабталған карталарды оперативты жаңартуда қолданылуы экономикалық тұрғыдан тиімді болып келеді. Үш спектралды каналды қашықтықтан зондтау негізінде құрылған түрлі-түсті суреттер жердік немесе аэрофотосуреттерге қарағанда көбірек ақпарат тасымалдайды, ал суреттердің стереожұптары кеңістік объектілерде үш өлшемді тасымалдау жүргізуге мүмкіндік береді [1,2].

Қазіргі таңда гиперспектральды суреттерде объектілерді айырып тануды жүзеге асыру алгоритмдері өте көп. Солардың ішіндегі көшбасшы алгоритмдердің бірі - k means кластерлеу алгоритмі.

Гиперспектральды суреттерде объектілерді айырып тануда k means кластерлеу алгоритмінің математикалық және сандық әдістерін моделдеу.

Java тілінде тізбектелген k means кластерлеу алгоритмі

Алгоритм итеративті түрде орындалады, ол берілген нүктелер(пиксель) жиынын олардың центрлеріне жақын келетін k кластер нүктелеріне бөліп, осы центрлердің орнының ауысуы нәтижесінде кластерлеу орындалады. k means алгоритмінде J

функциясы $J = \sum_{n=1}^N \sum_{k=1}^k \|x_n - c_k\|^2$ түрінде анықталады. Алгоритмнің мақсаты функцияны

минимумдау яғни функция қателігін квадраттау. $J = \sum_{n=1}^N \sum_{k=1}^k \|x_n - c_k\|^2$ - мұндағы, J

кластер центріне сәйкес N мәліметінің арасындағы ара-қашықтық, x_n ($1 \leq n \leq N$) берілген нүктелерді көрсетеді және c_k ($1 \leq k \leq K$) кластер ауырлығын анықтайды, $\|x_n - c_k\|^2$ - x_n

және c_k арасындағы ара-қашықтықты анықтайды. $\mu_k = \frac{1}{N_k} \sum_{n \in c_k} x_n$ - дағы, μ_k ($1 \leq k \leq K$)

берілген орташа нүкте мәні және N_k - k кластеріне жататын объектінің санын көрсетеді.

Біздің k means алгоритмінде, k кластер саны яғни қолданушы енгізген параметр. Екіншіден, кіру файлынан N объектілерін көрсетеді. Бастапқы жағдайда K центроидын кездейсоқ түрде таңдайды, μ_k ($1 \leq k \leq K$)-түрінде анықтайды. Екіншіден, алгоритм ара-қашықты есептеп, әрбір объектіге итеративті түрде кластерге жататын нүктелерді анықтайды. Қолданушының көрсеткен соңғы мәніне дейін процесс қайталанып отырады [3].

Java тілінде MPI Express кітапханасын қолдану арқылы параллельді кластерлеу алгоритмін шешу.

k means кластерлеу алгоритмі үлкен көлемді мәліметтер қорымен яғни жүз миллиондаған нүктемен және ондаған гигабайт мәліметтермен жұмыс істейді, осыған байланысты алгоритмді есептеуде параллельдеу өте тиімді.

Параллельді кластерлеу алгоритмі:

Input: K кластер саны, мәліметтер объектінің саны

Output: K центроиды

1: MPI_Init // процедураның басы;

2: N объектіні файлдан оқу;

3 Мәліметті берілген N объектіге процессор арасында бір қалыпта мәліметті бөледі және айталық әрбір процесс N объекті мәніне ие;

4: Әрбір процеске, 5-11 қадамын орындайды;

5: Әрбір объектіге меншіктейді x_n ($1 \leq n \leq N$) жақын кластеріне дейін;

6: $\mu_k = \frac{1}{N_k} \sum_{n \in c_k} x_n$ формуладан яғни, әрбір кластерден жаңа кластерді есептейді;

7: $J = \sum_{n=1}^N \sum_{k=1}^k \|x_n - c_k\|^2$ формуласынан J мәнін есептейді;

**ИНФОРМАТИКА. ИНФОРМАТИКАНЫ ОҚЫТУ ӘДІСТЕМЕСІ.
БІЛІМ БЕРУДІ АҚПАРАТТАНДЫРУ
ИНФОРМАТИКА, МЕТОДИКА ПРЕПОДАВАНИЯ ИНФОРМАТИКИ
ИНФОРМАТИЗАЦИЯ ОБРАЗОВАНИЯ**

8: Әрбір кластерден жаңа кластерді есептейді μ_k мына формуламен:

$$\mu_k = \frac{1}{N_k} \sum_{x_n \in C_k} x_n;$$

9: $J = \sum_{n=1}^N \sum_{k=1}^k \|x_n - c_k\|^2$ формуласынан J -ді есептейді;

10: 6-9 қадамы қайталанады, $J - J < \text{порог}$ ($J - J < \text{threshold}$) болмайынша;

11: Әрбір объект мәні үшін кластер идентификаторын құрады;

12: Әрбір итерация соңында, кластерлеу нәтижесіне байланысты жаңа центрлі есептейді;

13: Ең соңғы кластерлеу *Centroid* мәнін анықтайды;

14: MPI_Finalize // процедураның аяқталуы;

MapReduce негізінде k means кластерлеу алгоритмін шешу

MapReduce бұл бағдарламалық модель және пайдалы параллелизмді қамтамасыз етеді. Бағдарламалық модельде жазылған бағдарламалар автоматты түрде үлестірілген бағдарламада параллеленеді [4].

Қазіргі уақытта бейнелерді өте аз уақытта үлестірмелі өңдеу өзекті тақырып болып келеді. Ал MapReduce- Hadoop технологиясында үлестірмелі есептеуді қолдану қазіргі кездегі ғылыми-техника заманында негізгі тақырыптардың бірі.

Hadoop технологиясының негізгі артықшылығы - ол бірнеше компьютерлерге берілген кез келген үлкен көлемді деректерді үлестірмелі есептеуде автоматты түрде бөлін тастайды және масштабтайды, сонымен қатар Map және Reduce екі қадамы арқылы уақытты үнемдеп шығарып береді. Кез келген үлкен көлемді деректерді бір компьютерге есептеп, нәтижесін шешу қиындықтары болды. Бұл ғылыми қиындық Hadoop технологиясы пайда болғаннан кейін шешіле бастады. Күрделі деректерді үлестірмелі бейнелеуді өңдеуде қолданылады.

Hadoop технологиясында кластерлеу алгоритмі:

1. Бастапқыда, Mapper енгізілген мәндерді оқиды және бастапқы деректерлерді кіші деректер жиынына қысады, яғни қосалқы кластерге.

2. Әрбір Mapper, Reduce-ке жіберетін қосалқы кластерден k бастапқы кластер құрады.

3. Reduce әрбір Mapper-ден кластерді қосады және k кластер үшін центроидты қайта есептейді.

4. Осы центроидтар алмасу арқылы қайта бастапқы Mapper-ге жіберіледі.

5. Әрбір Mapper қосалқы кластерге арналған жаңа центроидты оқиды, Mapper кластерді қайта Reduce-ке жібереді.

6. Reduce кластерді қайта қосып жаңа центроидты есептейді.

7. Бұл процедура Mapper-ге мәліметті жіберуді Reduce тоқтатқанға дейін қайталанады. Бұл әдетте алгоритм теңескенде аяқталады.

Map функциясы:

(global object, in_key, in_value), global object contains the initial clustering centers, in_key has no usefulness, in_value is a string like (pixel_id, R, G, B). Output: (out_key, out_value), out_key is a string represents a clustering center, out_value is a same string as in_value.

1: construct initial clustering centers Array from global object;

2: labPixel = parseString (in_value);

3: minDistance = MAX_VALUE;

4: initial_array_subscript = -1;

```

5: for (j = 0; j < Array.length; ++j) {
6:   dist = cal_dist_labpixel_to_centers(labPixel, Array[j]); //бастапқы кластер
орталықтарын яғни енгізілген нүктелердің ара-қашықтықты есептейді.
   if (dist < minDistance) { minDistance = dist; initial_array_subscript = j; } }
7: out_key = Array[initial_array_subscript];
8: out_value = in_value; // пиксель орналасқан нүктені анықтайды
9: writeToHDFS(out_key, out_value);
10: output(out_key, out_value);
11: End;

```

Reduce функциясы:

Reduce function Input: (in_key, in_value), in_key is a string represents a clustering center, in_value is a string like (pixel_id, R, G, B).

Output: (out_key, out_value), out_key is a string represents the number of values which have the same key in iterator, out_value is a string represents a new clustering center after adjustment.

```

1: set the initial value of counter to 0;
2: set temp_ave like (null,null,null,null);
3: while(in_key.hasNext()) {
4:   temp_ave=temp_ave+abs(in_value.Next() - temp_ave)/(counter + 1);
   ++counter; } // жаңа орташа мәнді реттеп, итератор бос қалғанша жаңа орташа мәнді
есептейді.

```

```

5: out_key = counter.ToString();
6: out_value = temp_ave;
7: output(out_key, out_value);
8: End;

```

k -means кластерлеудің параллель алгоритмінің негізгі идеясы әр нүктенің(пиксель) жіктеуінде MapReduce негізде көрсетеді Map функцияда ең жақын кластер және жаңа кластерлердің орталарын есептеу барлық Reduce функциясында орындалады.

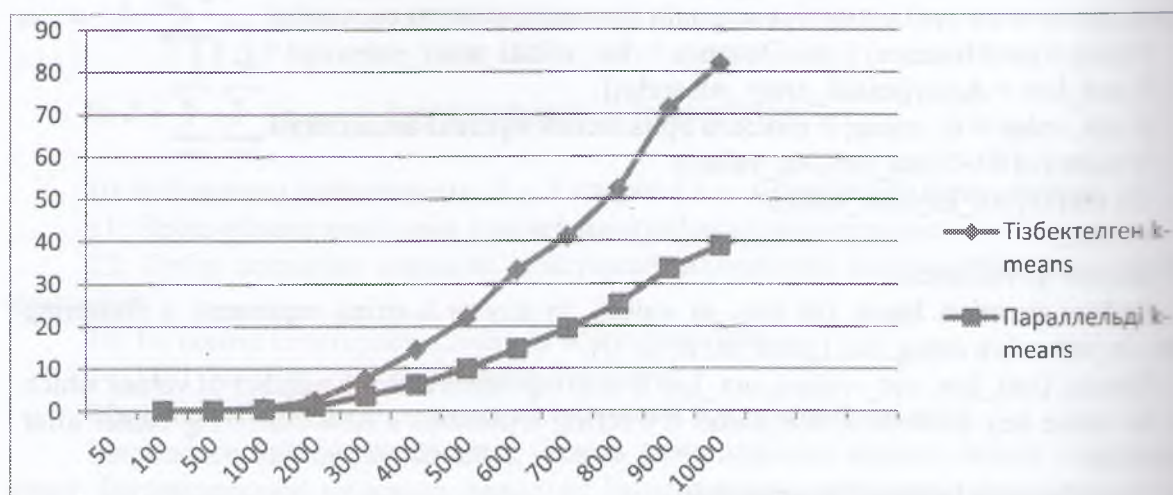
Сонымен, k means кластерлеу алгоритмінің тізбектелген және параллель нұсқасы Java бағдарламалау тілінде MPI технологиясын қолдану арқылы құрылды. Сонымен қатар, үлестірмелі алгоритмі MapReduce технологиясында жүзеге асырылды.

Кесте 1 - Тізбектелген және параллель k means кластерлеу алгоритмінің енгізілген нүкте мәніне байланысты өңдеу уақыты.

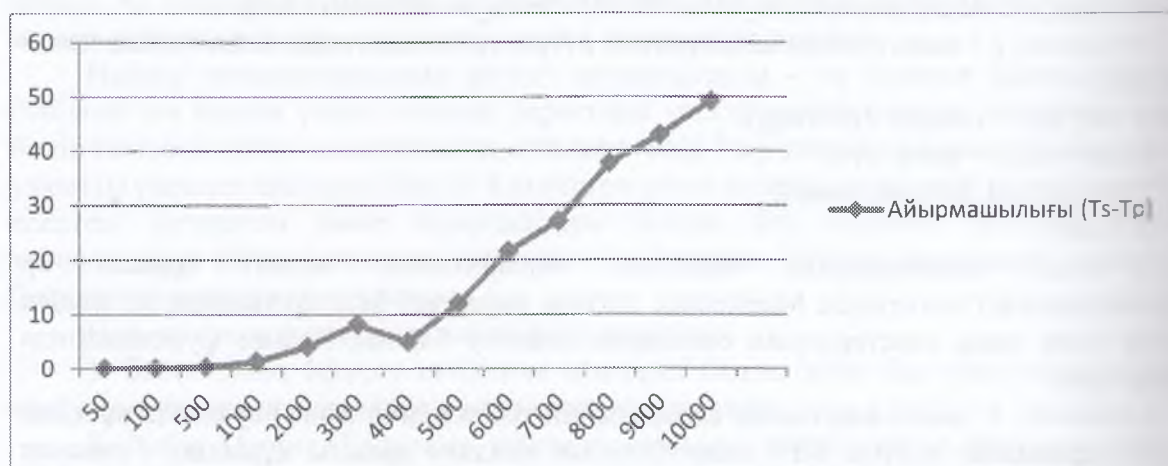
N нүкте мәні	Уақыт (Ts, sec) тізбектелген k means	Уақыт (Tp, sec) параллельді k means	Айырмашылығы (Ts-Tp)
50	0,087	0,032	0,055
100	0,175	0,0875	0,0875
500	0,6605	0,31025	0,31025
1000	2,462	1,131	1,301
2000	7,634	3,617	4,017
3000	14,345	6,1725	8,1725
4000	21,89	9,945	4,948
5000	33,155	14,5775	11,945
6000	41,06	19,51	21,55
7000	52,21	25,101	27,109
8000	71,3351	33,66755	37,66755
9000	81,91	38,955	42,955

**ИНФОРМАТИКА. ИНФОРМАТИКАНЫ ОҚЫТУ ӘДІСТЕМЕСІ.
БІЛІМ БЕРУДІ АҚПАРАТТАНДЫРУ
ИНФОРМАТИКА, МЕТОДИКА ПРЕПОДАВАНИЯ ИНФОРМАТИКИ
ИНФОРМАТИЗАЦИЯ ОБРАЗОВАНИЯ**

10000	92,02	42,862	49,158
-------	-------	--------	--------



Сурет 1 - Тізбектелген және параллельді кластерлеу алгоритмнің өңдеу уақыт аралығын салыстыру

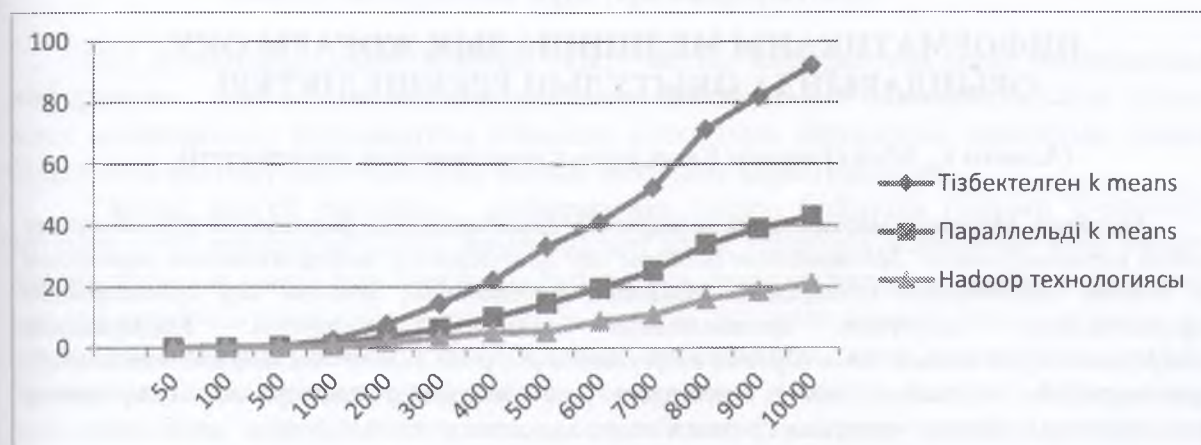


Сурет 2 - Тізбектелген және параллельді кластерлеу алгоритмнің өңдеу уақыт айырмашылығы

Кесте 2 - Java бағдарламалық ортасында MPJ Express кітапханасын қолдану арқылы және Hadoop технологиясында k means алгоритмнің енгізілген нүкте мәніне байланысты уақыт аралығы.

N нүкте мәні	Уақыт (Ts, sec) тізбектелген k means	Уақыт (Tp, sec) параллельді k means	Hadoop технологиясы
50	0,087	0,032	0,016
100	0,175	0,0875	0,04375
500	0,6605	0,31025	0,145125
1000	2,462	1,131	0,5655
2000	7,634	3,617	1,8085
3000	14,345	6,1725	3,08625
4000	21,89	9,945	4,9725
5000	33,155	14,5775	5,28875
6000	41,06	19,51	8,755
7000	52,21	25,101	10,83377
8000	71,3351	33,66755	15,83377

9000	81,91	38,955	18,4775
10000	92,02	42,862	20,928575



Сурет 3 - Тізбектелген, параллельді және үлестірілген кластерлеу алгоритмнің өңдеу уақыт аралығын салыстыру

Қорыта келе, MapReduce парадигмасын қолдану бейнені өңдеу кезіндегі алгоритмнің уақытын айтарлықтай қысқартты. Біздің тәжірбелерімізде есептеулер үлестірмелі есептеулерге арналған Hadoop платформасында, MapReduce парадигмасын қолданып есептелінді. Hadoop платформасын қолдану бізге есептің есептелуінің масштабын өзгертуге мүмкіндік берді, бірнеше есептеуіп түйіндерде есептелуін қамтамасыз етті.

Жасалған тәжірбиелерден үлестірмелі есептеулер есептеу жылдамдығын, түйіндер санын көбейту арқылы сызықты түрде жылдамдатуға қол жеткізуге болады. Тағыда жасалған тәжірбиелерден қарапайым компьютерлерді қолданып, Hadoop платформасында үлкен көлемді деректерді өңдеу тиімді екенін көреміз.

1. Кашкин В.Б., Сухинин А.И. Дистанционное зондирование Земли из космоса. Цифровая обработка изображений: Учебное пособие. – М.: Логос, 2001г. 264 стр.
2. В.В. Сергеев Анализ и обработка изображений, получаемых при наблюдениях земли из космоса// Стенограмма научного сообщения на совместном семинаре ИСОИ РАН и Института компьютерных исследований СГАУ 18 апреля 2006 года.
3. Р. Миллер, Л. Боксер. Последовательные и параллельные алгоритмы. Издательство Бином. Лаборатория знаний 2006г., 408стр.
4. J. Dean, S. Ghemawat. MapReduce: Simplified Data Processing on Large Clusters. Communications of The ACM, 2008. 51(1), 107-113.

Аннотация. Один из наиболее популярных методов кластеризации является алгоритм *k means*, из-за его легкой реализации, простоте, эффективности и эмпирических успехов. Целью данного исследования является проведение экспериментов алгоритма *k means* в технологии Hadoop и реализовать параллельный алгоритм на языке Java с обращениями к библиотеке MPJ Express.

Ключевые слова: кластеризация, оптимизация параллельных вычислений, распределенная программа.

Abstract. The most famous clustering algorithm is *k means* because of its easy implementation, simplicity, efficiency and empirical success. The goal of this study is to perform *k means* clustering using Hadoop and implement a parallel algorithm in Java with calls library MPJ Express.

Keywords: clustering, optimization of parallel computing, distributed program.