КОМПЬЮТЕРНАЯ МАТЕМАТИКА

УДК 519.854.33

В.П. ШИЛО, В.А. РОЩИН, И.П. ГРАДИНАР

ПРИБЛИЖЕННОЕ РЕШЕНИЕ ЗАДАЧИ НАХОЖДЕНИЯ МАКСИМАЛЬНОГО ρ-ПЛОТНОГО МНОЖЕСТВА ВЕРШИН ГРАФА

Предложены постановка и приближенный алгоритм решения задачи нахождения максимального ρ -плотного множества вершин графа. Изучены свойства такого множества, приведены результаты экспериментальных расчетов.

Введение. Задача нахождения максимального множества вершин графа с допустимой плотностью имеет практическое применение в различных сетях (социальных, биологических, экономических) и в других приложениях [1]. В известных задачах нахождения максимального независимого множества или максимальной клики [2], максимального k -plex или co- k -plex [3] присутствует ограничение на число вершин, связанных с каждой вершиной. В отличие от них в исследуемой задаче вводится ограничение на отношение общего числа ребер графа к числу его вершин, т. е. нами рассматривается более обобщенное понятие связности графа. Подобным задачам посвящены работы [1, 4]. Особенно близка к рассматриваемой задача нахождения максимальной квази-клики [1]. Отличие квази-клики от ρ -плотного множества состоит в том, что плотность последнего ограничена сверху. Ясно, что необходимое для квази-клики условие связности в рассматриваемой нами задаче можно опустить.

В данной работе дается определение ρ -плотного множества вершин графа, приводится постановка задачи нахождения такого множества с максимальным числом вершин, предлагается алгоритм ее решения.

Постановка задачи. Пусть задан граф $G(V,E)$ с множеством вершин V и множеством ребер E .

Плотностью графа $G(V,E)$ называется величина

$$D(G)=\frac{2|E|}{|V|(|V|-1)},$$

где $|V|$ – мощность множества V.

Очевидно, что понятие плотности следует рассматривать при $|V|\geq 2$.

Предположим, что $G[V']=(V',E\cap(V'\times V'))$ – подграф графа $G(V,E)$, порожденный множеством $V'\subset V$.

Множество $V'\subset V$ вершин графа $G(V,E)$ (и соответственно порожденный им граф $G[V']$) назовем ρ-плотным ($0\leq\rho\leq 1$), если для него выполняется условие

$$D(G[V'])\leq\rho. \qquad (1)$$

Задача нахождения максимального ρ-плотного множества вершин графа $G(V,E)$ заключается в отыскании множества $V'\subset V$, имеющего максимально возможное число вершин и удовлетворяющего условию (1).

Взаимосвязь с другими задачами. Рассматриваемая задача тесно связана с задачей поиска максимального множества $V'\subset V$ вершин графа $G(V,E)$, для которого выполняется условие

$$D\left(G[V']\right)\geq\rho',\ \ 0\leq\rho'\leq 1.$$

Это видно из следующих утверждений.

Утверждение 1. Пусть V' – ρ-плотное множество вершин графа $G(V,E)$. Тогда оно является множеством вершин графа $\overline{G(V,E)}$, для которого справедливо неравенство

$$D\left(\overline{G[V']}\right)\geq 1-\rho, \qquad (2)$$

где $\overline{G(V,E)}$ – дополнение графа $G(V,E)$.

Доказательство. Поскольку V' – ρ-плотное множество вершин графа $G(V,E)$, то для него выполняется условие (1). Тогда

$$D\left(\overline{G[V']}\right)=\frac{2\left|E\left(\overline{G[V']}\right)\right|}{|V'|(|V'|-1)}=\frac{2\left(\frac{|V'|(|V'|-1)}{2}-\left|E\left(G[V']\right)\right|\right)}{|V'|(|V'|-1)}=$$

$$=\frac{|V'|(|V'|-1)-2\left|E\left(G[V']\right)\right|}{|V'|(|V'|-1)}=1-\frac{2\left|E\left(G[V']\right)\right|}{|V'|(|V'|-1)}=1-D\left(G[V']\right)\geq 1-\rho.$$

Утверждение доказано.

Утверждение 2. Если V' – максимальное ρ-плотное множество вершин графа $G(V,E)$, то оно – максимальное множество вершин графа $\overline{G(V,E)}$, удовлетворяющее условию (2).

Доказательство. Так как V' – ρ-плотное множество вершин графа $G(V,E)$, то в соответствии с утверждением 1 оно является множеством вершин графа $\overline{G(V,E)}$, для которого выполняется условие (2). Остается доказать, что V' – максимальное множество вершин графа $\overline{G(V,E)}$, удовлетворяющее неравенству (2). Допустим противное: имеется множество $V'' \subset V$, $|V''| > |V'|$, которое удовлетворяет условию (2). Тогда

$$D(G[V'']) = \frac{2\left|E\left(G[V'']\right)\right|}{|V''|(|V''|-1)} = 1 - \left(1 - \frac{2\left|E\left(G[V'']\right)\right|}{|V''|(|V''|-1)}\right) =$$

$$= 1 - \frac{|V''|(|V''|-1) - 2\left|E\left(G[V'']\right)\right|}{|V''|(|V''|-1)} = 1 - \frac{2\left(\frac{|V''|(|V''|-1)}{2} - \left|E\left(G[V'']\right)\right|\right)}{|V''|(|V''|-1)} =$$

$$= 1 - \frac{2\left|E\left(\overline{G[V'']}\right)\right|}{|V''|(|V''|-1)} = 1 - D\left(\overline{G[V'']}\right) \le 1 - (1-\rho) = \rho.$$

Таким образом, $D(G[V'']) \le \rho$, а это противоречит условию утверждения. Утверждение доказано.

Свойства ρ-плотных множеств. Представим свойства ρ-плотных множеств вершин графа в виде следующих утверждений.

Теорема. Пусть задан граф $G(V,E)$, $V = \{v_1,...,v_n\}$, с плотностью $D(G)$, тогда существует по крайней мере одна его вершина v_i, $i \in I = \{1,...,n\}$, такая, что $D(G[V \setminus v_i]) \le D(G)$.

Доказательство. Предположим, что $D(G) = \rho$ и для всех $i \in I$ $D(G[V \setminus v_i]) > \rho$. Обозначим $m = |E|$, $m_i = |E(G[V \setminus v_i])|$, $\rho_i = D(G[V \setminus v_i])$, $i \in I$. Число вершин графа $G[V \setminus v_i]$,, $i \in I$, равно $n-1$. Если в графе $G(V,E)$ имеется ребро между вершинами v_j и v_k, то оно присутствует и в графе $G[V \setminus v_i]$, $i \in (I \setminus \{j,k\})$. Тогда

$$\sum_{i \in I} m_i = (n-2)m .$$

$$\rho_i = D\left(G[V \setminus v_i]\right) = \frac{2m_i}{(n-1)(n-2)} > \rho, \ i \in I.$$

Просуммируем ρ_i

$$\sum_{i\in I}\rho_i = \frac{2\sum_{i\in I} m_i}{(n-1)(n-2)} = \frac{2(n-2)m}{(n-1)(n-2)} = \frac{2m}{n-1}.$$

Тогда

$$D(G) = \frac{2m}{n(n-1)} = \frac{\sum_{i\in I}\rho_i}{n} > \frac{n\rho}{n} = \rho.$$

Таким образом, приходим к противоречию с условием теоремы. Теорема доказана.

Легко видеть, что из данной теоремы вытекает справедливость таких следствий.

Следствие 1. Пусть задан граф $G(V,E)$, $V=\{v_1,...,v_n\}$, с плотностью $D(G)\le\rho$, тогда существует последовательность различных индексов $i_1,\ldots,i_n$, $i_j\in I$, $j\in I$, таких, что множества $\{v_{i_1},v_{i_2}\},\ldots,\{v_{i_1},\ldots,v_{i_{n-1}}\}$ будут также ρ-плотными.

Следствие 2. Если для графа $G(V,E)$, $V=\{v_1,...,v_n\}$, выполняется условие $D(G)\le\rho$, то существует такая последовательность всех вершин из V, что множества $V\setminus\{v_{i_1}\}$, …, $V\setminus\{v_{i_1},\ldots,v_{i_{n-2}}\}$, $i_j\in I$, $j\in I$, будут ρ-плотными.

На рисунке схематично отображена суть теоремы и следствий 1, 2.

Описание алгоритма. Пусть $N_G(v)$ – множество вершин неориентированного графа $G(V,E)$, инцидентных вершине $v\in V$, т. е.

$$N_G(v)=\{u\in V \mid \exists e=(u,v)\in E\}.$$

Обозначим $deg_G(v)$ степень вершины v (число вершин, инцидентных вершине v), т. е. $deg_G(v)=|N_G(v)|$.

Предлагаемый в данной работе приближенный алгоритм базируется на использовании вышеприведенной теоремы и следствий 1, 2. На их основе при добавлении вершин в искомое максимальное ρ-плотное множество V' графа $G(V,E)$ или удалении вершин из него должно выполняться условие (1), т. е. множество V' всегда ρ-плотно. Отсюда следует, что в V' может быть занесена такая вершина v из $V\setminus V'$, для которой выполняется неравенство

$$\frac{2(|E(G[V'])|+\deg_{G[V'\cup v]}(v))}{|V'|(|V'|+1)}\le\rho, \tag{3}$$

а из V' может быть удалена вершина v, удовлетворяющая условию

$$\frac{2(|E(G[V'])|-\deg_{G[V']}(v))}{(|V'|-1)(|V'|-2)}\le\rho. \tag{4}$$

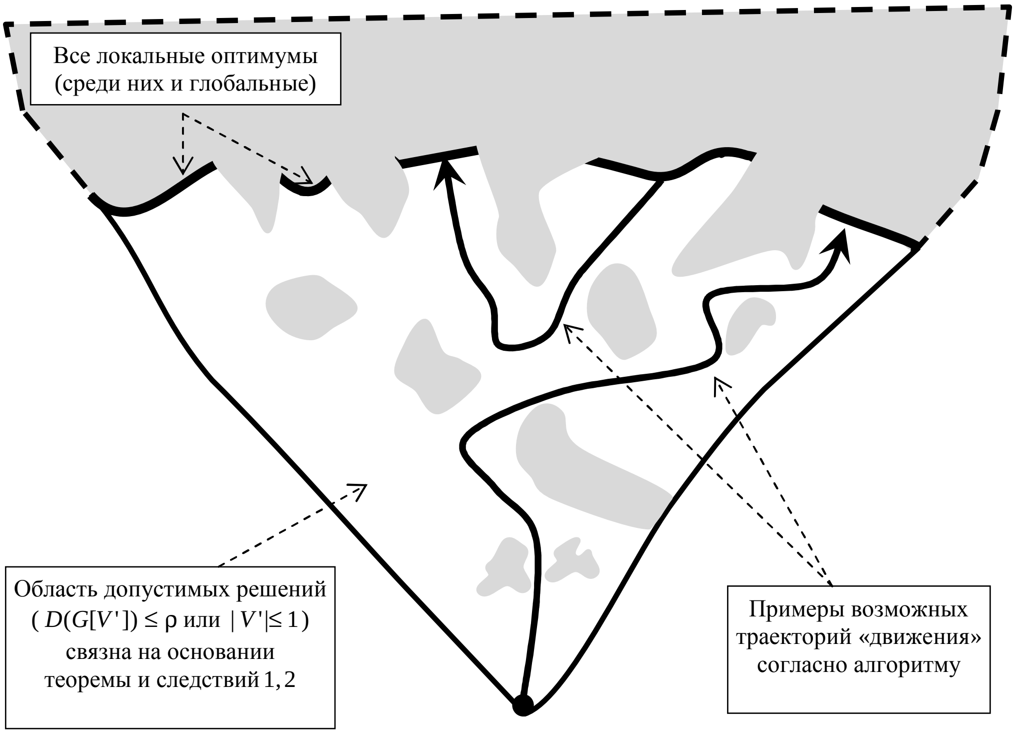

РИСУНОК. Схематичная иллюстрация теоремы и следствий 1, 2

Вначале работы алгоритма полагаем $V' = \varnothing$. Далее осуществляем случайный поочередный набор в множество V' вершин из $V \setminus V'$, удовлетворяющих условию (3) и имеющих минимальное значение $\deg_{G[V' \cup v]}(v)$, до момента достижения локального оптимума. После этого случайным образом удаляем из V' вершины, для которых выполняется неравенство (4), и опять, по тому же правилу, наполняем множество V'. Все эти действия повторяем до выполнения некоторого заданного критерия останова (например, времени, выделенного на решение задачи). Следует отметить, что иногда (с малой вероятностью) можно осуществлять случайный выбор вершин из множества $V \setminus V'$ не только среди тех, для которых значение $\deg_{G[V' \cup v]}(v)$ минимально, но и среди всех вершин, удовлетворяющих условию (3).

Общую схему алгоритма можно представить в таком виде:

1. *largest_cardinality* = 0;
2. **while** *global_criterion* ( )
3. $\quad V' = \varnothing$;

```
        filling_V' ( );
        if largest_cardinality<|V'|
                largest_cardinality = |V'|;
        end if
        while local_criterion ( )
                i = choosing_value ( );
                deleting_some_verticies_from_V' (i);
                filling_V' ( );
                if largest_cardinality<|V'|
                        largest_cardinality = |V'|;
                end if
        end while
end while
return largest_cardinality.
```

Результаты экспериментальных расчетов. Данный алгоритм был использован для решения задачи нахождения максимального ρ-плотного множества вершин дополнений DIMACS-графов [5]. Программная реализация предложенного алгоритма была скомпилирована с помощью C++ и запущена (при загрузке 25%) на компьютере с конфигурацией: Intel® Core™2 Quad CPU, Q9550 @ 2.83GHz, 8,00GB RAM. Значение ρ для экспериментальных расчетов выбиралось следующим образом: плотность каждого графа умножалась на коэффициент l, который принимал значения от 0 до 0,8 с шагом 0,2 ($\rho_l = D(G) * l$). Каждая задача решалась 10 раз с ограничением по времени в 1 минуту для конкретного ρ. Полученные результаты представлены в таблице. Здесь приведено число вершин и ребер, а также плотность графа. В колонке *a* содержится наибольшее значение мощности найденных 10-ти максимальных ρ-плотных множеств; в колонке b – среднее значение полученных максимальных мощностей; в колонке *c* – среднее время (в сек.), затраченное на получение наилучшего результата при каждом запуске алгоритма.

Заключение. В работе сформулирована постановка задачи нахождения максимального ρ-плотного множества вершин графа, изучены свойства такого множества. Предложен также алгоритм решения рассматриваемой задачи, приведены полученные для тестовых графов максимальные мощности ρ-плотных множеств при различных значениях ρ.

ТАБЛИЦА. Результаты вычислительных экспериментов

$\overline{G(V,E)}$	$\lvert V\rvert$	$\lvert E\rvert$	$D(G)$	Коэффициент l ($\rho_l = D(G)\cdot l$)														
				0			0,2			0,4			0,6			0,8		
				Полученные результаты														
				a	b	c	a	b	c	a	b	c	a	b	c	a	b	c
brock200-1	200	5066	0,2546	21	21	0,52	32	32	0,02	46	46	0,44	71	71	0,57	119	119	0,06
brock200-2	200	10024	0,5037	12	12	0,15	16	16	≤0,01	24	24	≤0,01	40	40	0,02	79	79	≤0,01
brock200-3	200	7852	0,3946	15	15	0,34	21	21	0,03	31	31	≤0,01	52	52	0,09	96	96	0,03
brock200-4	200	6811	0,3423	17	17	0,76	24	24	≤0,01	36	36	≤0,01	57	57	0,03	103	103	≤0,01
brock400-1	400	20077	0,2516	27	25,4	7,9	39	39	0,17	61	61	5,4	102	102	0,31	193	193	0,18
brock400-4	400	20035	0,2511	33	33	5,9	39	39	0,1	61	61	0,14	103	103	4,3	191	191	0,49
brock800-1	800	112095	0,3507	21	20,9	20	35	35	16	57	57	2,4	106	106	2,8	238	238	5,6
brock800-2	800	111434	0,3487	21	21	21	35	35	6	57	57	8	105	104,9	13	240	240	6
brock800-4	800	111957	0,3503	21	21	24	34	34	3,7	57	57	7,7	102	102	6,1	233	232,6	12
c-fat500-1	500	120291	0,9643	14	14	≤0,01	18	18	≤0,01	27	27	≤0,01	42	42	0,07	83	83	0,16
c-fat500-2	500	115611	0,9267	26	26	≤0,01	33	33	≤0,01	49	49	0,04	75	75	0,08	137	137	0,16
c-fat500-5	500	101559	0,8141	64	64	0,3	80	80	0,12	111	111	0,45	160	160	1,8	244	244	1,9
c-fat500-10	500	78123	0,6262	126	126	0,84	147	147	0,72	195	195	2,6	251	251	0,86	324	324	1,4
hamming6-2	64	192	0,0952	32	32	≤0,01	33	33	≤0,01	35	35	≤0,01	39	39	≤0,01	44	44	≤0,01
hamming6-4	64	1312	0,6508	4	4	≤0,01	4	4	≤0,01	6	6	≤0,01	13	13	≤0,01	34	34	≤0,01
hamming8-2	256	1024	0,0314	128	128	≤0,01	135	135	≤0,01	144	144	≤0,01	156	156	≤0,01	176	176	≤0,01
hamming8-4	256	11776	0,3608	16	16	≤0,01	19	19	≤0,01	34	34	≤0,01	112	112	≤0,01	150	150	≤0,01
hamming10-2	1024	5120	0,0098	512	512	0,32	540	540	0,04	576	573,3	6,9	627	627	0,06	707	707	≤0,01
hamming10-4	1024	89600	0,1711	40	40	2,2	65	65	9,4	149	148,4	12	537	537	≤0,01	626	626	≤0,01
johnson32-2-4	496	14880	0,1212	16	16	≤0,01	17	17	≤0,01	21	21	≤0,01	34	34	≤0,01	66	66	≤0,01
keller4	171	5100	0,3509	11	11	≤0,01	16	16	≤0,01	29	29	≤0,01	59	59	≤0,01	105	105	≤0,01
keller5	776	74710	0,2485	27	27	1,5	48	48	5,8	132	132	1,7	288	284,5	8,1	482	482	1,8
keller6	3361	1026582	0,1818	55	54,1	38	159	155,6	34	592	583,2	19	1273	1235,4	9,7	2065	2055,2	0,32
MANN-a9	45	72	0,0727	16	16	≤0,01	18	18	≤0,01	22	22	≤0,01	28	28	≤0,01	36	36	≤0,01
MANN-a27	378	702	0,0099	124	123,7	23	144	143,1	8,5	214	204,2	24	352	352	≤0,01	364	364	≤0,01
MANN-a45	1035	1980	0,0037	335	334	16	393	390,3	16	609	593,3	24	994	994	≤0,01	1014	1014	≤0,01
MANN-a81	3321	6480	0,0012	1084	1082,1	0,04	1276	1272	0,03	2179	2178,4	0,06	3252	3252	0,09	3285	3285	0,1
p-hat700-2	700	122922	0,5024	44	44	0,04	162	162	≤0,01	290	290	≤0,01	432	432	≤0,01	566	566	≤0,01
p-hat700-3	700	61640	0,2520	62	62	0,19	176	176	≤0,01	298	298	0,02	434	434	≤0,01	567	567	≤0,01
p-hat1000-1	1000	377247	0,7552	10	10	0,04	18	18	0,03	36	36	0,04	97	97	0,9	419	419	0,06
p-hat1000-2	1000	254701	0,5099	46	46	0,21	196	196	≤0,01	393	393	0,03	596	596	≤0,01	797	797	≤0,01
p-hat1000-3	1000	127754	0,2558	68	68	3,2	214	214	≤0,01	402	402	0,02	604	604	0,02	801	801	≤0,01
p-hat1500-1	1500	839327	0,7466	12	11,6	3,8	21	21	≤0,01	42	42	0,24	127	127	0,21	643	643	0,08
p-hat1500-2	1500	555290	0,4939	65	65	0,42	335	335	0,04	635	635	0,05	926	926	0,02	1210	1210	0,02
p-hat1500-3	1500	277006	0,2464	94	94	5,5	347	347	0,98	641	641	0,08	931	931	0,07	1213	1213	0,02
san200-0.7-1	200	5970	0,3000	30	30	0,13	105	105	0,05	114	114	0,08	126	126	0,04	146	146	0,04
san200-0.7-2	200	5970	0,3000	18	18	0,7	43	43	0,02	141	141	≤0,01	155	155	≤0,01	172	172	≤0,01
san200-0.9-1	200	1990	0,1000	70	70	0,08	128	128	≤0,01	140	140	≤0,01	155	155	≤0,01	175	175	≤0,01
san200-0.9-2	200	1990	0,1000	60	60	≤0,01	105	105	≤0,01	118	118	0,02	135	135	≤0,01	160	160	0,04
san200-0.9-3	200	1990	0,1000	44	44	0,03	59	59	0,18	106	106	0,02	123	123	0,06	151	151	0,06
san400-0.5-1	400	39900	0,5000	13	13	0,71	25	25	0,17	224	224	0,89	248	248	0,27	285	285	0,74
san400-0.7-1	400	23940	0,3000	40	40	5,6	203	203	≤0,01	220	220	≤0,01	244	244	1,3	281	281	1,3
san400-0.7-2	400	23940	0,3000	30	30	14	138	138	≤0,01	221	221	1,1	245	245	1,0	283	283	0,69
san400-0.7-3	400	23940	0,3000	22	22	11	51	51	6,5	229	229	0,26	254	254	0,25	294	294	0,97
san400-0.9-1	400	7980	0,1000	100	100	0,83	205	205	0,67	225	225	0,96	253	253	0,91	298	298	1,4
san1000	1000	249000	0,4985	15	11,5	5,9	43	43	2,4	562	554,8	20	618	594,3	8,5	708	708	14
sanr200-0.7	200	6032	0,3031	18	18	0,05	28	28	≤0,01	41	41	0,03	63	63	≤0,01	110	110	≤0,01
sanr200-0.9	200	2037	0,1024	42	42	0,25	58	58	≤0,01	80	80	0,03	109	109	≤0,01	154	154	≤0,01
sanr400-0.5	400	39816	0,4989	13	13	0,7	20	20	0,13	32	32	0,25	57	57	0,55	126	126	0,23
sanr400-0.7	400	23931	0,2999	21	21	0,64	34	34	0,27	53	53	0,24	91	91	0,29	183	183	0,15

В.П. Шило, В.О. Рощин, І.П. Градинар

НАБЛИЖЕНЕ РОЗВ'ЯЗАННЯ ЗАДАЧІ
ЗНАХОДЖЕННЯ МАКСИМАЛЬНОЇ ρ -ЩІЛЬНОЇ МНОЖИНИ ВЕРШИН ГРАФА

Запропоновано постановку та наближений алгоритм розв'язання задачі знаходження максимальної ρ -щільної множини вершин графа. Вивчено властивості такої множини, наведено результати експериментальних розрахунків.

V.P. Shylo, V.O. Roshchyn, I.P. Gradinar

APPROXIMATE ALGORITHM FOR THE MAXIMUM ρ -DENSE SET PROBLEM

In the paper, a formulation and approximate algorithm for solving the maximum ρ -dense set problem is proposed. The properties of such a set are studied and the results of computer experiments are presented.

1. *Abello J., Resende M., and Sudarsky S.* Massive Quasi-Clique Detection // In Proceedings of Latinoamerican Informatics (LATIN 2002, May). – Berlin: Springer, 2002. – P. 598–612. – Available at: http://www.mgvis.com/Papers/MassiveDataSets/massive_quasiclique.pdf (Last visited: November 22, 2010).
2. *Сергиенко И.В., Шило В.П.* Задачи дискретной оптимизации: проблемы, методы решения, исследования. – Киев: Наук. думка, 2003. – 264 с.
3. *Шило В.П., Градинар І.П., Ляшко В.І.* Наближений алгоритм знаходження максимального k -plex (co- k -plex) графа // Наукові записки НаУКМА: Комп'ютерні науки. – 2011. – **112**.
4. *Brunato M., Hoos H.H., Battiti R.* On Effectively Finding Maximal Quasi-Cliques in Graphs // Proc. 2nd Learning and Intelligent Optimization Workshop (LION2007 II, Trento, Italy, December 2007). LNCS 5313 – Berlin: Springer, 2008. – P. 41–55. – Available at: http://dit.unitn.it/~brunato/pubblicazioni/lionII-quasi.pdf (Last visited: November 30, 2010).
5. *DIMACS*: Maximum clique, graph coloring, and satisfiability. Second DIMACS implementation challenge – Available at: http://dimacs.rutgers.edu/Challenges/ (Last visited: 05 December 2010).

Получено 10.12.2010

Об авторах:

Шило Владимир Петрович,
доктор физико-математических наук, ведущий научный сотрудник
Института кибернетики имени В.М. Глушкова НАН Украины,
e-mail: v.shylo@gmail.com

Рощин Валентина Алексеевна,
кандидат физико-математических наук, старший научный сотрудник
Института кибернетики имени В.М. Глушкова НАН Украины,
e-mail: d135@i.com.ua

Градинар Иван Петрович,
аспирант Ужгородского национального университета.
e-mail: vgradinar@gmail.com