

Т.В. Селівьорстова

ДОСЛІДЖЕННЯ ПОВЕДІНКИ МУЛЬТИАГЕНТНОЇ СИСТЕМИ НА ОСНОВІ ІМІТАЦІЙНОЇ МОДЕЛІ СТРАТЕГІЧНОЇ ПОВЕДІНКИ

Анотація. Розглянуті постановки задачі поведінки рухомих об'єктів в рамках класу задач, який носить умовну назву «задача про дифузійну бомбу». Виконано програмну реалізацію та дослідження моделі поведінки рухомих об'єктів, яка полягає в «кооперативному» децентралізованому прийнятті рухомими об'єктами рішень по вибору траєкторій руху, в рамках заданої цілі.

Ключові слова: Стратегічна поведінка, задача про дифузійну бомбу, кооперативна модель, рефлексивно-кооперативна модель.

Постановка проблеми. У багатьох прикладних областях виникають завдання управління групою рухомих об'єктів (РО), що спільно виконують деяке завдання, наприклад, пошук рухомих або нерухомих об'єктів в заданій області простору, або в проникненні в задану область, або в ураженні цілей та ін. [1]. Як правило, група РО функціонує в конфліктному середовищі, тобто в умовах протидії (виявлення, інформаційної протидії, знищення) – з боку об'єктів пошуку, системи охорони або оборони (елементи якої умовно називають «сенсорами»), що забезпечує захист кордонів області, цілей та ін.

Аналіз публікацій по темі дослідження. В ході дослідження даної проблематики [2 – 6] склалась система класифікацій задач групового керування в умовах протидії, яка передбачає варіативність кількості та якості цілей, ступеня ураження, часового проміжку, кількості РО, швидкості РО, мережі сенсорів та їхньої кількості, взаємодії РО та сенсорів, що утворило розгалужену класифікацію задач, зокрема:

– планування траєкторій руху (Route Planning, Cooperative Path Planning) множини РО в відсутності протидії (в неконфліктній середовищі) [3];

– дискретне планування траєкторій руху з урахуванням протидії (Optimal Risk Path Planning) [4];

– розподілу ЦО (Coordinated Target Assignment) [5, 6].

Формулювання цілей статті. Враховуючи, що однією з важливих постановок є задача «пошуку» (ураження) на площині нерухомого ЦО без часових обмежень при наявності кількох нерухомих сенсорів (РО рухаються з постійною по абсолютній величині заданій швидкістю (напрямок руху може змінюватися); планування ними траєкторій здійснюється децентралізовано (автономно) в реальному часі). У якості критерію ефективності дій групи РО вважатимемо число K РО, які досягли ЦО. Даний клас задач носить умовну назву «задача про дифузійну бомбу». Цілю дослідження є реалізація та дослідження моделі поведінки РО, яка полягає в «кооперативному» децентралізованому прийнятті рухомими об'єктами рішень по вибору траєкторій руху в умовах, коли ймовірність виявлення (знищення) кожного з них залежить від відносного розташування всіх членів їх групи.

Планування траєкторії в умовах протидії

Задані початкові положення $(x_j(0), y_j(0))$, $j = \overline{1, K_0}$, на площині K_0 рухомих об'єктів. Їх мета – з'явитися в точці з координатами (x^*, y^*) . Положення j -го РО в момент часу $t \geq 0$ позначимо через $(x_j(t), y_j(t))$, його швидкість – через $v_j(t) = \sqrt{(\dot{x}_j)^2 + (\dot{y}_j)^2}$, час першого потрапляння в точку (x^*, y^*) – через T_j . Є N нерухомих сенсорів з координатами (a_i, b_i) , $i = \overline{1, N}$, мають можливість підсумувати приходять на них в один і той же момент часу сигнали. Відстань від j -го РО до i -го сенсора позначимо через $\rho_{ij}(t) = \sqrt{(x_j(t) - a_i)^2 + (y_j(t) - b_i)^2}$.

У загальному випадку ризик виявлення j -го РО системою сенсорів описується наступним функціоналом:

$$R_j = \int_0^{T_j} \sum_{i=1}^N \frac{(v_j(t))^m}{(\rho_{ij}(t))^k} dt, \quad (1)$$

де «сигнал» на сенсорі (доданок у виразі (1)) залежить від швидкості РО і відстані від останнього до сенсора. З вигляду функціоналу (1) випливає, що ризик виявлення РО залежить від значень «сигналів»

на різних сенсорах. Показник ступеня k характеризує фізичне поле, в якому здійснюється виявлення [7], а показник ступеня m – залежність рівня інтенсивності випромінюваного сигналу від швидкості руху об'єкта (наприклад, сигналів первинного гідроакустичного поля [8]). Значення k відповідають:

- $k = 1$ – первинному гідроакустичному полю в мілкому морі;
- $k = 2$ – тепловому полю, первинного електромагнітного поля

і первинного гідроакустичного поля в глибокому морі;

- $k = 3$ – магнітному полю;
- $k = 4$ – вторинному (при активному режимі виявлення) електромагнітному і гідроакустичному полям.

«Некооперативна» модель

Нехай всі РО рухаються з постійною по абсолютній величині швидкістю v_0 . Знаючи розташування сенсорів і їх умовні невід'ємні чутливості $\{c_i\}$, $i = \overline{1, N}$, можливо за аналогією з виразом (1) для кожної точки (x, y) площини визначити ризик (ймовірність виявлення)

$$r(x, y) = \min \left\{ \sum_{i=1}^N \frac{c_i}{\left(\sqrt{(x - a_i)^2 + (y - b_i)^2} \right)^k}; 1 \right\} \quad (2)$$

виявлення окремого РО, що знаходиться в цій точці.

Нехай час дискретний. Крок часу позначимо через τ , через p позначимо ймовірність знищення виявленого РО (для простоти будемо вважати, що ця ймовірність не залежить від координат точки виявлення, часу і швидкості РО – облік в майбутніх дослідженнях цих залежностей є перспективним), через

$e(x, y) = (x^* - x, y^* - y) / \sqrt{(x - x^*)^2 + (y - y^*)^2}$ – одиничний вектор напрямку на ЦО в точці (x, y) , через $\rho((x, y); (q, w))$ – евклідова відстань між точками (x, y) і (q, w) , через $s_\Delta(x, y)$ – коло радіуса $\Delta \geq 0$ з центром в точці (x, y) .

Розглянемо кілька стратегій поведінки РО.

Варіант I. Перший (найпростіший) варіант, коли кожен РО рухається по прямій, що з'єднує його початкове положення з ЦО. Відповідний РО умовно назвемо неінтелектуальним.

В рамках варіанту I кожен РО в кожен момент часу повинен знати тільки своє поточне становище і становище ЦО.

Більш «інтелектуальні» РО повинні враховувати поточні та (або) майбутні ймовірності їх виявлення. Для опису їх поведінки визначимо безліч таких точок, що:

1) в них РО може виявитися, почавши рухатися з точки (x, y) зі швидкістю v_0 , через час τ ;

2) ймовірність виявлення РО не перевищує граничної величини δ ;

3) плюс сама точка (x, y) :

$$S_{v_0\tau}^\delta = \{(q, w) \mid \rho((x, y); (q, w)) = v_0\tau; r(q, w) \leq \delta\} \cup \{(x, y)\}. \quad (3)$$

Лінію рівня сумарного сигналу $r(q, w) = \delta$ будемо називати пороговою лінією. Позначимо $\text{Pr oj}_{S_{v_0\tau}^\delta(x, y)}(x^*, y^*)$ – проекцію положення ЦО на $S_{v_0\tau}^\delta(x, y)$ (т.ч. найближча в сенсі евклідова відстань до ЦО точка цієї множини; якщо така проекція не єдина, виберемо рівномірно будь-яку з них).

Варіант II. Введемо наступні правила планування РО своїх траєкторій (алгоритм некооперативного поведінки).

Крок 1. У кожен період часу РО, що знаходиться в точці (x, y) , з ймовірністю $p r(x, y)$ знищується системою оборони, а з ймовірністю $1 - p r(x, y)$ продовжує рух.

Крок 2. Продовживши рух, РО до початку наступного періоду часу виявиться в точці (u, v) , де

$$(u, v) \in \begin{cases} \text{Pr oj}_{S_{v_0\tau}^\delta(x, y)}(x^*, y^*), & \text{якщо } (x, y) \notin \text{Pr oj}_{S_{v_0\tau}^\delta(x, y)}(x^*, y^*), \\ \text{інакше } (x, y) + \min\{v_0\tau; \rho((x, y); (x^*, y^*))\} e(x, y). \end{cases} \quad (4)$$

Перший випадок у виразі (1.4) відповідає невіддаленню від ЦО за умови, що ризик виявлення не перевищить пороговий. Другий випадок – «прорив» по прямій до ЦО (в разі, якщо не вдається забезпечити виконання умови неперевикнення ймовірністю виявлення порогового значення, у випадку коли проходження цій умові призводить до віддалення від ЦО).

В рамках алгоритму некооперативного поведінки, РО, що знаходиться в деякій точці, повинен мати оцінки ризику виявлення тільки для $v_0\tau$ -околиці цієї точки; таким чином поведінка РО локально-оптимальна, і вимагає тільки локальної інформації. Відзначимо, що при активному режимі виявлення ($k = 4$) ця локальна інформація може обчислюватися за допомогою екстраполяції поточних вимірювань сигналів сенсорів. При пасивному режимі досить зажадати знання усіма РО координат і чутливості сенсорів (що дає їм можливість розрахувати ризик (2) для довільної точки площини).

Некооперативна поведінка групи РО буде описуватися наступним чином: для кожного моменту часу для кожного РО виконувати послідовно крок 1 і крок 2 до тих пір, поки або не будуть знищені всі РО, або всі вцілілі РО не досягнуть ЦО.

«Кооперативна» модель

Взаємодію РО будемо враховувати в такий спосіб: припустимо, що ймовірність виявлення даного РО залежить не тільки від поточних відстаней від нього до сенсорів, а й від того, наскільки близько до нього розташовані інші РО (приклад – зростання ефективної поверхні розсіювання). Іншими словами, умовно можна вважати, що РО служать «сенсорами» один для одного, і в міру їх взаємного зближення зростає ймовірність виявлення.

Позначимо через

$$R_j(x_j, y_j) = \min \left\{ r(x_j, y_j) + \sum_{l \neq j} \frac{\alpha}{1 + \left(\sqrt{(x_j - x_l)^2 + (y_j - y_l)^2} \right)^k}; 1 \right\} \quad (5)$$

ризик виявлення j -го РО, що знаходиться в точці (x_j, y_j) , з урахуванням його взаємодії з іншими РО, де α – невід’ємна константа.

Варіант III. Рухливі об’єкти прориваються до ЦО по прямій, не враховуючи і не прогножуючи ймовірності їх виявлення. Даний варіант відповідає варіанту I з точністю до заміни ризику (2) на ризик (5). Інформованість РО при цьому така ж, що і у варіанті I.

Варіант IV. Алгоритм «кооперативного» поведінки буде описуватися кроками 1' і 2', які з точністю до заміни ризику (2) на ризик (5) збігаються відповідно з кроками 1 та 2, причому у виразі (5)

підсумовування ведеться по тим РО, які до поточного моменту не були знищені.

Варіант IV відповідає варіанту II з точністю до заміни ризику (2) на ризик (5).

В даному випадку для планування своєї траєкторії кожен РО, крім інформації, необхідної в другому варіанті, повинен знати поточні координати всіх РО. (Можливо узагальнення «кооперативної» моделі на випадок, коли кожен РО має свій фіксований «радіус огляду» і при плануванні своєї траєкторії має інформацію і враховує (у натуральному вираженні типу (5)) тільки ті інші РО, які знаходяться від нього на відстані, що не перевищує цей радіус.)

Рефлексивна модель

Будемо вважати, що в групі присутні РО двох типів. Об'єкти першого типу – назвемо їх нерексуючими – діють відповідно до алгоритму «кооперативного» поведінки (варіант IV). Об'єкти другого типу – назвемо їх рефлексуючі [14] – діють більш складним чином: кожен з них, вважаючи всіх інших нерексуючими [15], прогнозує їх поведінку. Іншими словами, той РО, що рефлексує, точно розраховує, де опиняться в наступний момент часу інші РО (діють відповідно до варіанту IV) і вибирає напрямок свого руху з урахуванням прогнозованих положень інших РО.

Визначимо кроки 1'' і 2'' як кроки, які з точністю до заміни ризику (5) на прогнозований ризик збігаються відповідно з кроками 1' і 2'.

Варіант V. Алгоритм рефлексивної поведінки групи РО: для кожного моменту часу для кожного нерексуючого РО виконувати послідовно кроки 1' і 2', а для кожного рефлексуючого РО виконувати послідовно кроки 1'' і 2'', поки всі вцілілі РО не досягнуть ЦО.

Відзначимо, що в рамках алгоритму рефлексивної поведінки (варіант V) інформованість кожного РО повинна бути такою ж, що і в разі кооперативного поведінки (варіант IV).

Реалізація імітаційних моделей стратегічної поведінки

Програмна реалізація та візуалізація результатів роботи імітаційних моделей стратегічної поведінки проводилась у середовищі MatLab з використанням встроєної мови програмування.

Вхідними параметрами некооперативної моделі руху РО є:

– кількість агентів, для кожного агента задаються його координати та стан (0 – агент живий, 1 – агент знищений, 2 – агент досяг ЦО);

– кількість сенсорів, для кожного сенсору задається його координати та стан (0 – безпечний, 1 – небезпечний);

– розмір обчислювального поля;

– межа виявлення (границя детектування) РО.

Критерієм проведення обчислень є існування агентів з міткою 0 – рухливі незнищенні або ті, що не потратили до цільової точки.

В результаті проведення обчислювального експерименту, що реалізує некооперативну модель руху РО при граничному значення $\delta = 0,3$ для трьох сенсорів (1, 2, 5) з небезпечним сигналом, було встановлено, що жоден з агентів не дістався точки цілі. На рисунку 1 наведена динаміка руху РО в процесі досягнення цілі для 100 агентів (червоний колір), які розташовані у 5 рядів по 20 з відстанню 2 кроки дискретизації, 7 сенсорів, що розташовані по кутах латинської букви W. Сенсори, що безпечні пофарбовані зеленим кольором, а небезпечні – жовтим кольором. Об'єкт цілі – синя точка.

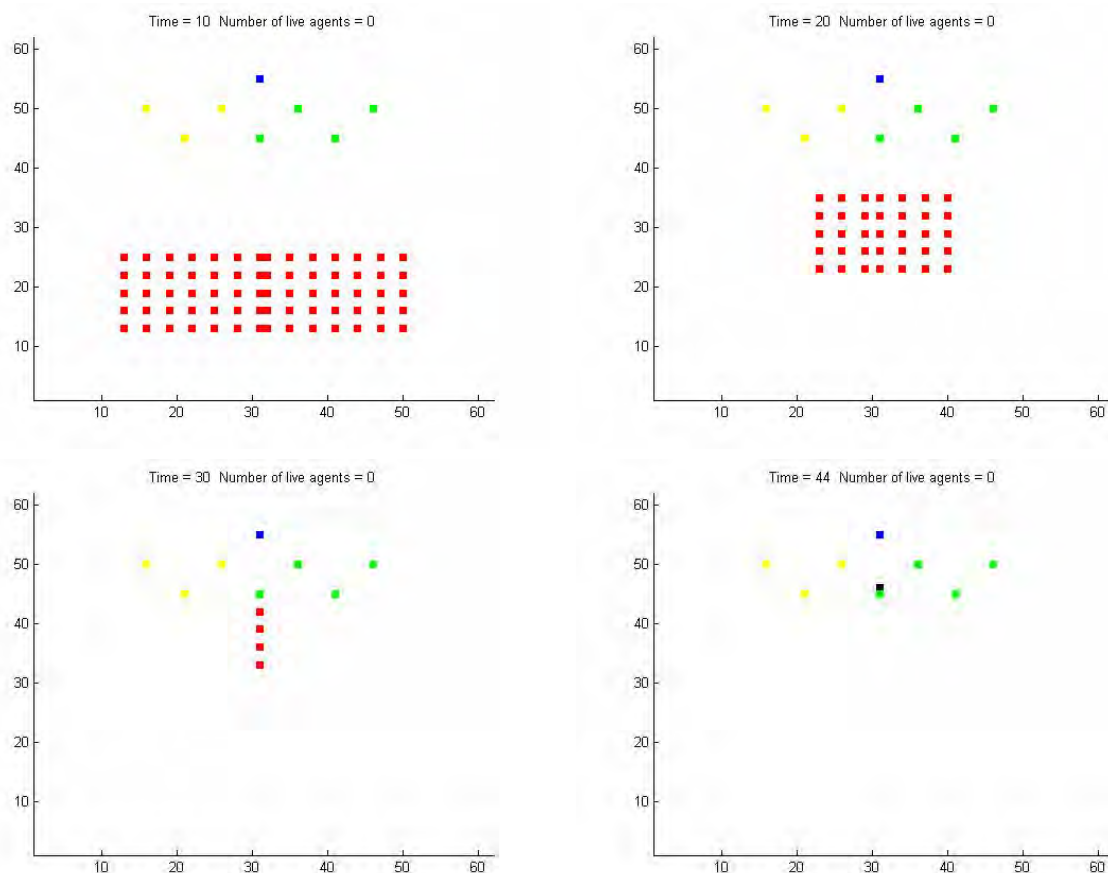


Рисунок 1 – Візуалізація результату некооперативної моделі

Кооперативна модель відрізняється від некооперативної тим, що скупчення агентів збільшує ризик бути знищеним, тобто сенсорами небезпеки для рухомих об'єктів є у тому числі інші рухомі об'єкти. Вводиться додатковий коефіцієнт α , який визначає чутливість агентів до присутності інших РО. Динаміка руху РО згідно кооперативної моделі при $\delta = 0,7$, коефіцієнті кооперації $\alpha = 0,01$ для чотирьох сенсорів (1, 2, 3) з небезпечним сигналом, зображена на рисунку 2, відображає досягнення цілі 76 агентами.

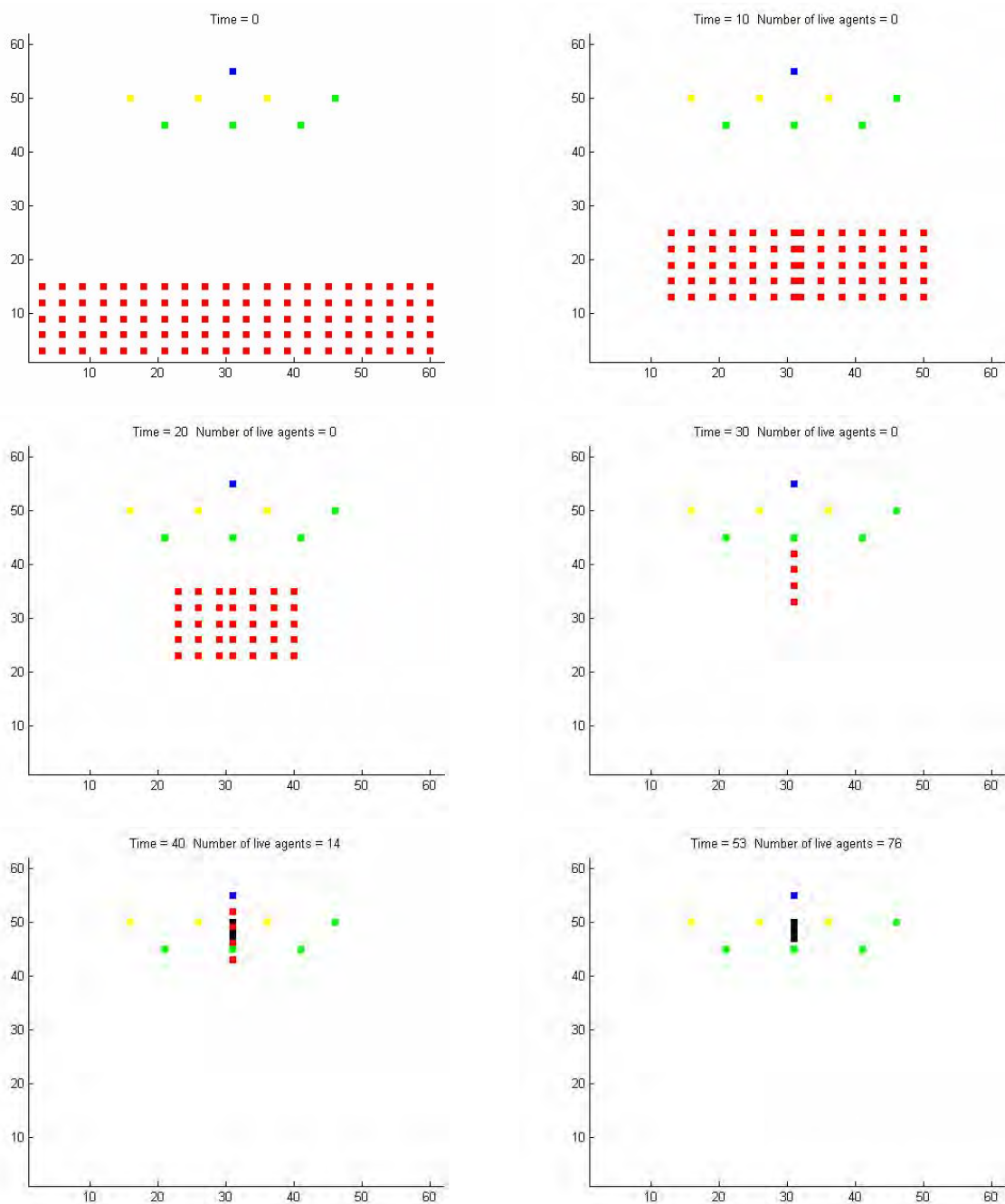


Рисунок 2 – Візуалізація результату кооперативної моделі

Рефлексивно-кооперативна модель, описує рух об'єктів при врахуванні ризику знищення, який обчислюється з врахуванням міркувань про те, що сенсорами до рухомих об'єктів є у тому числі інші рухомі об'єкти.

Модель працює наступним чином: проводиться прогноз напрямку руху, обчислюється ризик у точці прогнозованого руху, якщо ризик знищення менше граничного значення – агент рухається, у протилежному випадку агент або зостається на місці, або переходить у будь-яку точку околу, незважаючи на ризик знищення. Відмінності від інших реалізацій полягають у випадковому переході у точки околу, який реалізується за допомогою генератора псевдовипадкових чисел.

На рисунку 3 відображено результат обчислень за рефлексивно-кооперативною моделлю при граничному значенні $\delta = 0,5$, коефіцієнті кооперації $\alpha = 0,01$ для трьох сенсорів (1, 2, 3) з небезпечним сигналом. Цілі досягли 54 агента.

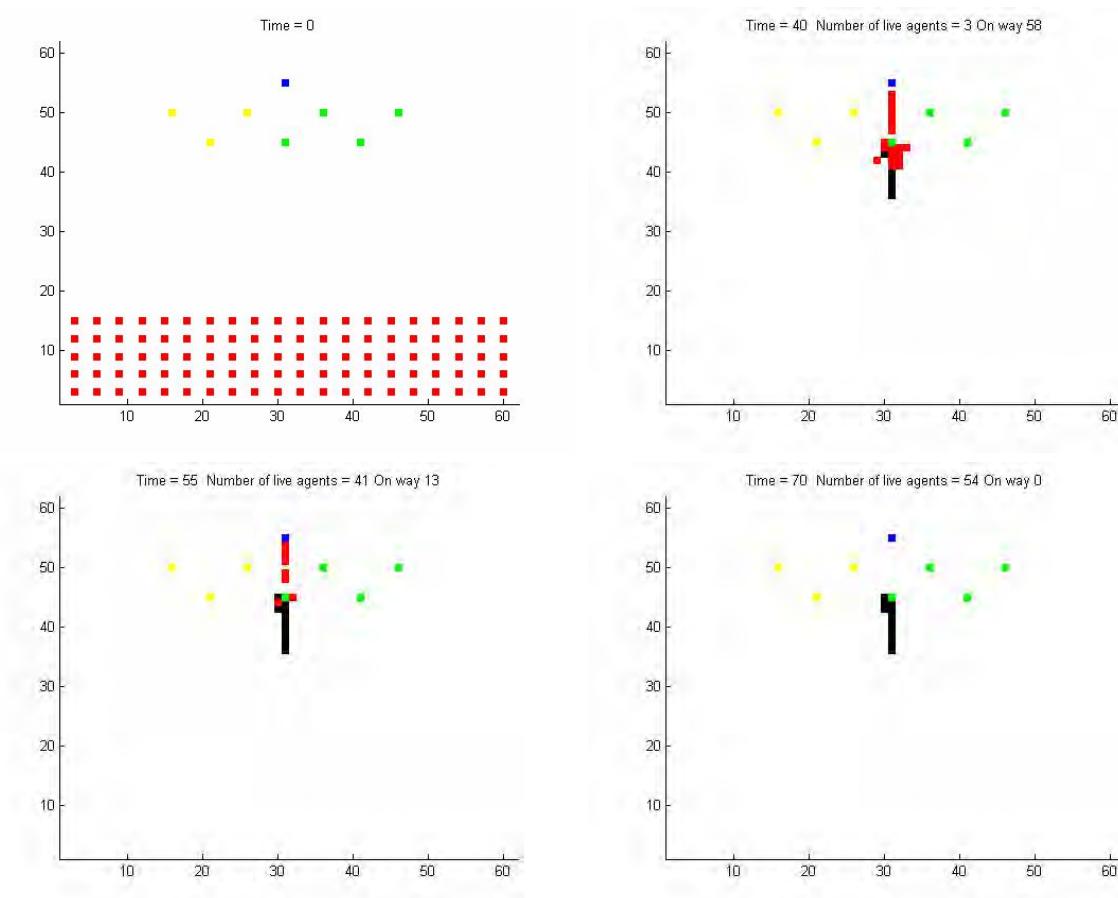


Рисунок 3 – Візуалізація результату рефлексивно-кооперативної моделі

**Дослідження роботи моделей стратегічної поведінки у
мультиагентних системах**

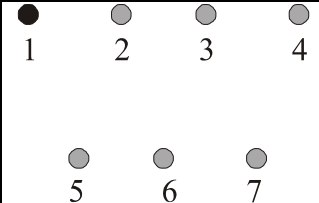
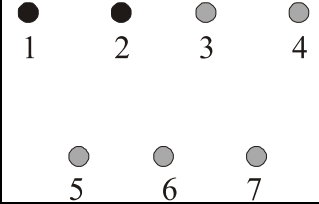
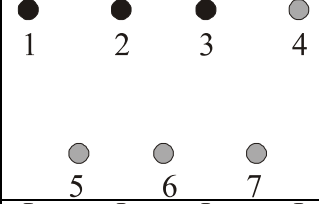
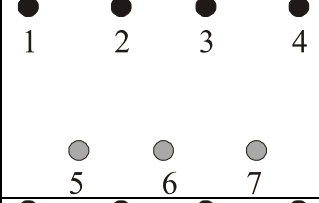
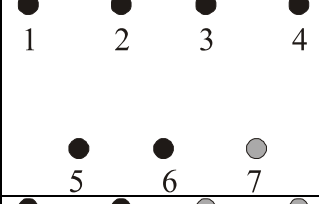
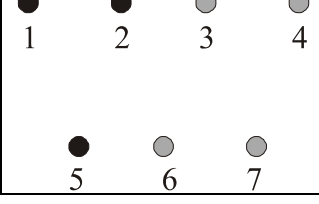
Завдяки реалізованим моделям було проведено низку обчислювальних експериментів, результати яких представлені в таблицях 1 – 3.

Таблиця 1

Результати моделювання отримані нереклексивною моделлю

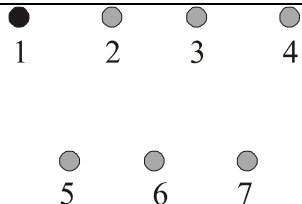
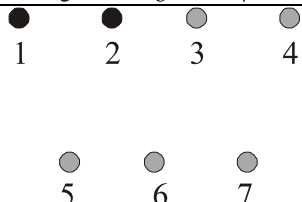
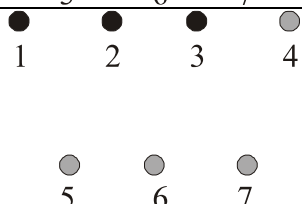
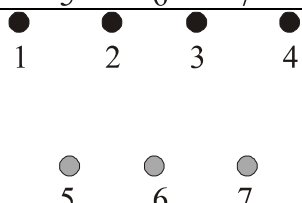
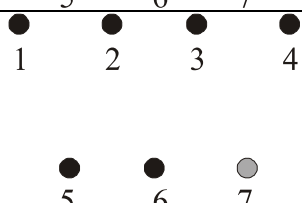
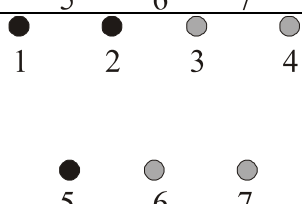
Розташування сенсорів	Кількість агентів, що досягли цільової точки		
	$\delta = 0,3$	$\delta = 0,5$	$\delta = 0,7$
	100	100	100
	100	100	100
	0	100	100
	0	0	100
	0	0	0
	0	100	100

Результати моделювання отримані кооперативною моделлю

Розташування сенсорів	Кількість агентів, що досягли цільової точки					
	$\alpha = 0,01$			$\alpha = 0,03$		
	$\delta = 0,3$	$\delta = 0,5$	$\delta = 0,7$	$\delta = 0,3$	$\delta = 0,5$	$\delta = 0,7$
	75	100	100	26	46	67
	7	76	100	3	26	48
	0	5	76	0	2	24
	0	0	49	0	0	17
	0	0	0	0	0	0
	0	44	100	0	16	37

В таблиці 3 наведені усереднені дані по результатах 3-х обчислювальних експериментів, у зв'язку з тим, що в алгоритмі присутній випадковий чинник.

Результати моделювання отримані
рефлексивно-кооперативної моделлю

Розташування сен- сорів	Кількість агентів, що досягли цільової точки								
	$\alpha = 0,00$			$\alpha = 0,01$			$\alpha = 0,03$		
	0,3	0,5	0,7	0,3	0,5	0,7	0,3	0,5	0,7
 1 2 3 4 5 6 7	100	100	100	67	99	100	30	50	69
 1 2 3 4 5 6 7	100	100	100	12	85	100	6	34	53
 1 2 3 4 5 6 7	0	100	100	0	10	94	0	4	41
 1 2 3 4 5 6 7	0	0	100	0	0	73	0	0	28
 1 2 3 4 5 6 7	0	0	0	0	0	0	0	0	0
 1 2 3 4 5 6 7	0	100	100	0	64	96	0	22	42

На базі проведених обчислювальних експериментів були зроблені наступні висновки:

1) для некооперативної моделі – при зниженні порогу детектування кількість агентів, що досягли цільової точки зменшується,

причому для даної моделі властиво те, що або всі агенти досягають цільової точки, або всі знищуються системою безпеки;

2) для кооперативної моделі безпеки – при зменшенні коефіцієнту кооперативного детектування більше агентів досягає цільової точки, також при зниженні порогу детектування кількість агентів, що досягли цільової точки зменшується (рисунок 4);

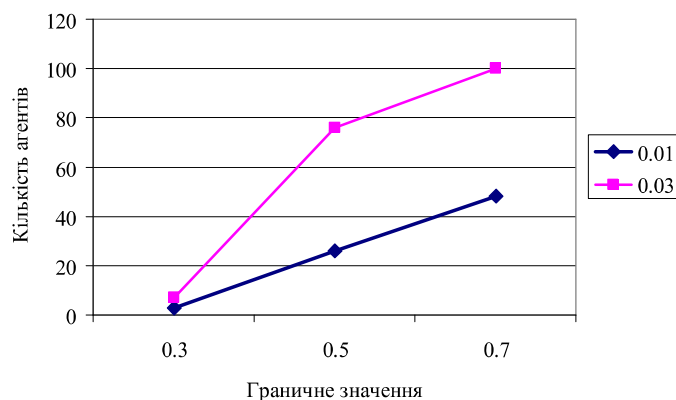


Рисунок 4 – Залежність кількості агентів, що досягли цілі від граничного значення при варіації параметру кооперації, для активних сенсорів 1, 2

3) для рефлексивно-кооперативної моделі – завдяки рефлексивній компоненті кількість агентів, що досягли цільової точки збільшується, при зменшенні коефіцієнту кооперативного детектування більше агентів досягає цільової точки, також при зниженні порогу детектування кількість агентів, що досягли цільової точки зменшується (рисунок 5, 6).

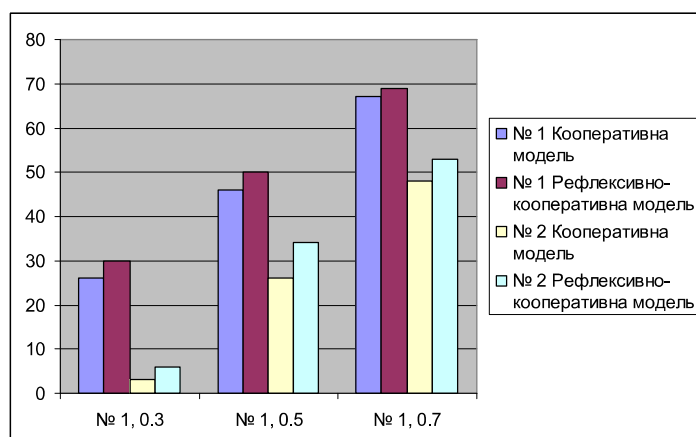


Рисунок 5 – Залежність кількості агентів, що досягли цілі від граничного значення при варіації моделей, для обчислювального експерименту №1 та № 2

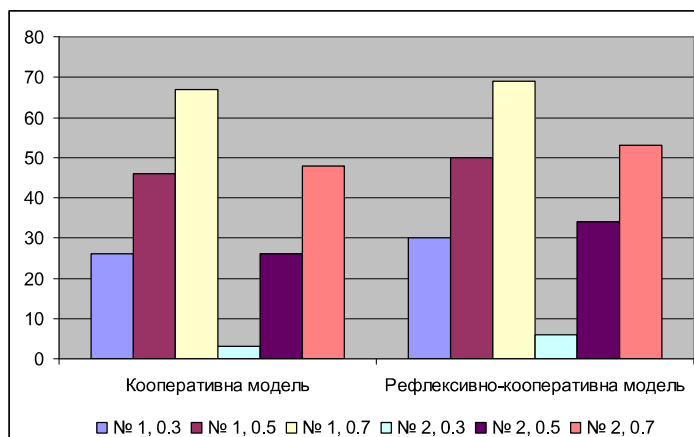


Рисунок 6 – Залежність кількості агентів, що досягли цілі при варіації моделей для обчислювальних експериментів №1 та № 2

ЛІТЕРАТУРА

1. Каляев И.А., Гайдук А.Р., Капустян С.Г. Модели и алгоритмы коллективного управления в группах роботов. – М.: Физматлит, 2009.
2. Liu B., Dousse O., Wang J., Saipulla A. Strong Barrier Coverage of Wireless Sensor Networks / ACM MobiHoc08, 2008.
3. Zheng C., Li L., Xu F, Sun F. Evolutionary Route Planner for Unmanned Air Vehicles // IEEE Transactions on Robotics. – 2005. – Vol. 21, № 4. – P. 609 – 620.
4. Anisi D., Thunberg J. Swedish Defense Research Agency, 2007. Report FOI-R-2266-SE. – 68 p.
5. Добровидов А.В., Кулида Е.Л., Рудько И.М. Выбор траектории движения объекта в конфликтной среде // Проблемы управления. – 2011. – № 2. – С. 64 – 75.
6. Beard R., McLainy T., Goodrich M. Coordinated Target Assignment and Intercept for Unmanned Air Vehicles // Proceedings of IEEE2002 International Conference on Robotics and Automation. Washington DC, 2002. – P. 2581 – 2586.
7. Абрамянц Т.Г., Маслов Е.П., Яхно В.П. Уклонение подвижного объекта от обнаружения группой наблюдателей // Проблемы управления. – 2010. – № 5. – С. 73 – 79.
8. Абрамянц Т.Г., Маслов Е.П., Рудько И.М., Яхно В.П. Уклонение подвижного объекта от обнаружения группой наблюдателей при малых отношениях сигнал/помеха // Обработка информации и управление. – 2011. – № 2.
9. Zabarankin M., Uryasev S., Pardalos P. Optimal Risk Path Algorithms // Cooperative Control and Optimizatон. Ch. 1 / Eds. Murphey R., Pardalos P. – Dordrecht: Kluwer Acad., 2002. – P. 271 – 303.