

УДК 528.29

О. В. Боровик,*доктор технічних наук, професор, начальник навчального відділу Національної академії Державної прикордонної служби України імені Богдана Хмельницького,***Р. В. Рачок,***кандидат технічних наук, доцент, докторант докторантури Національної академії Державної прикордонної служби України імені Богдана Хмельницького,***В. Ю. Мазур,***кандидат військових наук, доцент, докторант докторантури Національної академії Державної прикордонної служби України імені Богдана Хмельницького*

Методика виявлення просторових аномалій руху суден та її використання при оцінюванні ризиків у системі морської охорони «Гарт-12»

У статті представлена методика виявлення просторових аномалій руху суден на основі даних в інформаційно-телекомунікаційній системі морської охорони «Гарт-12». У межах цієї методики запропоновано: метрику для визначення подібності двох маршрутів суден на основі дискретних відліків про їхнє положення в різні моменти часу; метод кластеризації для обробки наявних у «Гарт-12» даних і визначення типових маршрутів. При оцінюванні ризиків порушення прикордонного законодавства на морській ділянці кордону запропоновано враховувати відповідність просторових даних маршруту його описовій частині, середнє та максимальне відхилення від типового маршруту. При перевірці описових даних використане нечітке порівняння текстової інформації на основі метрики Левенштейна.

Ключові слова: інформаційно-телекомунікаційна система морської охорони «Гарт-12», система висвітлення надводної обстановки, оцінювання ризиків, кластеризація, метрика, маршрут.

© О. В. Боровик, Р. В. Рачок, В. Ю. Мазур, 2018

Постановка проблеми в загальному вигляді. Забезпечення національної безпеки України потребує ефективної охорони державного кордону. Значною та важливою частиною останнього є морська ділянка. З метою її ефективної охорони в Державній прикордонній службі України (ДПСУ) використовується система постів технічного спостереження, інформація з яких надходить до інформаційно-телекомунікаційної системи (ІТС) морської охорони «Гарт-12». Важливим елементом ІТС «Гарт-12» є система висвітлення надводної обстановки (СВНО), яка забезпечує відображення даних про положення суден і допоміжної інформації щодо їхньої назви, портів відправлення та призначення тощо.

Використання СВНО пов'язане з роботою відповідних експертів (операторів). А отже, ефективність застосування СВНО значною мірою залежить від низки суб'єктивних факторів, зокрема підготовленості оператора, його психофізичного стану тощо. Необхідність опрацювання значної кількості даних у системі ще більше ускладнює оцінювання обстановки при прийнятті рішень.

Це пояснює той факт, що задача підвищення ефективності використання ІТС «Гарт-12» є актуальною та її розв'язання може залежати від реалізації в ній елементів підтримки прийняття рішень загалом і реалізації механізму оцінювання ризиків порушення прикордонного законодавства на морській ділянці кордону на основі даних про надводну обстановку зокрема.

Аналіз останніх досліджень і публікацій, у яких започатковане розв'язання даної проблеми та на які спираються автори. Необхідність імплементації до інформаційно-телекомунікаційних систем ДПСУ елементів підтримки прийняття рішень, зокрема з використанням геообробки даних, наведена в дослідженнях [1–2]. У роботі [2] визначена доцільність проведення аналізу даних надводної обстановки з метою виявлення просторово-часових аномалій руху суден, однак не наведені шляхи вирішення цього завдання.

У дослідженні [3] приділена значна увага розв'язанню задачі виявлення просторових аномалій руху суден на основі аналізу даних у СВНО. У межах цієї роботи були визначені можливі підходи до просторового аналізу даних в ІТС «Гарт-12». Із цією метою була розроблена метрика для порівняння окремих маршрутів на основі дискретних даних про них, отриманих у різні моменти часу. Ця метрика дає змогу визначити ступінь подібності маршрутів. Однак, як буде показано далі, вона не є кумулятивною, що вимагає її вдосконалення. Не вирішеними в дослідженні [3] залишилися і питання класифікації даних про маршрути та оцінювання ризиків.

Особливості представлення даних в ІТС «Гарт-12» обмежують можливість використання класичних методів кластеризації [4] для їх класифікації та визначення типових маршрутів руху суден.

Для аналізу ризиків у складних системах відомим є застосування низки методів, одним з яких є SWOT-аналіз

[5]. Однак специфіка просторового аналізу даних надводної обстановки в деяких випадках обмежує його можливість. Тому виявлення небезпек правопорушень у прикордонній сфері на морській ділянці кордону потребує опрацювання загальних підходів, які не мали б указаних недоліків.

Мета статті полягає в розробці методики виявлення просторових аномалій руху суден та її використання при оцінюванні ризиків у ІТС морської охорони «Гарт-12».

Виклад основного матеріалу дослідження. Дані про рух суден у базі геоданих ІТС «Гарт-12» представляють собою впорядковані послідовності точок з координатами суден, отриманими в певні моменти часу. Позначимо сукупність таких координат для i -го маршруту впорядкованою множиною (кортежем) $M_i = \{(x_1, y_1), \dots, (x_{k_i}, y_{k_i})\}$, де k_i – потужність множини M_i .

Для оцінювання ризиків порушення прикордонного законодавства на морській ділянці кордону на основі виявлення аномалій руху суден вбачається за необхідне розв'язання низки часткових задач:

- проведення класифікації наявних даних з визначенням типових маршрутів руху суден;
- аналіз нових даних щодо їх належності до одного з типових маршрутів та оцінювання ступеня просторових відхилень;
- визначення ступеня ризику порушення прикордонного законодавства на морській ділянці кордону.

Важливим елементом розв'язання наведених вище задач є встановлення ступеня подібності маршрутів, який можна оцінити на основі метрики, введеної в роботі [2]. Оскільки окремі відліки отримані в різні моменти часу, то для порівняння маршрутів пропонується проводити відновлення одного з маршрутів на основі апроксимації та знаходити середнє відхилення від нього дискретних точок іншого маршруту. Тоді метрику $R(M_1, M_2)$ для маршрутів, які описуються двома множинами точок M_1 та M_2 , можна визначити на основі [2]

$$R(M_1, M_2) = \frac{\sum_{j=1}^{k_2} \min_i (R_{\min_{ij}})}{k_2}, \quad (1)$$

де $R_{\min_{ij}}$ – найменша відстань від i -го сегмента апроксимованого маршруту M_1 , який поєднує точки (x_i, y_i) і (x_{i+1}, y_{i+1}) множини M_1 , до j -ї точки множини M_2 .

Оскільки в базі даних ІТС «Гарт-12» маршрути звичайно описуються значною кількістю відліків, можна використати лінійну апроксимацію. Тоді параметр $R_{\min_{ij}}$ визначається з виразу

$$R_{\min_{ij}} = \begin{cases} \frac{|(y_{i+1} - y_i)x_j + (x_i - x_{i+1})y_j + (y_i x_{i+1} - x_i y_{i+1})|}{\sqrt{(y_{i+1} - y_i)^2 + (x_i - x_{i+1})^2}}, & x_{ij}^n \in [x_i, x_{i+1}], \\ \min \left\{ \sqrt{(x_j - x_i)^2 + (y_j - y_i)^2}, \sqrt{(x_j - x_{i+1})^2 + (y_j - y_{i+1})^2} \right\}, & x_{ij}^n \notin [x_i, x_{i+1}]. \end{cases} \quad (2)$$

де (x_i, y_i) , (x_{i+1}, y_{i+1}) – координати i -ї та $(i+1)$ -ї точок множини M_1 ;

(x_j, y_j) – координати j -ї точки множини M_2 .

x_{ij}^n – координата (по осі x) проекції j -ї точки множини M_2 на пряму, яка проходить через i -й сегмент.

Однак при такому визначенні метрики $R(M_1, M_2)$ остання може відрізнятися від $R(M_2, M_1)$. Тобто метрика (1) не є комутативною. Це пояснюється тим, що можливі випадки, коли точки другого маршруту є близькими до апроксимованого першого маршруту, але точки першого маршруту від апроксимованого другого маршруту розташовані на більших відстанях (рис. 1).

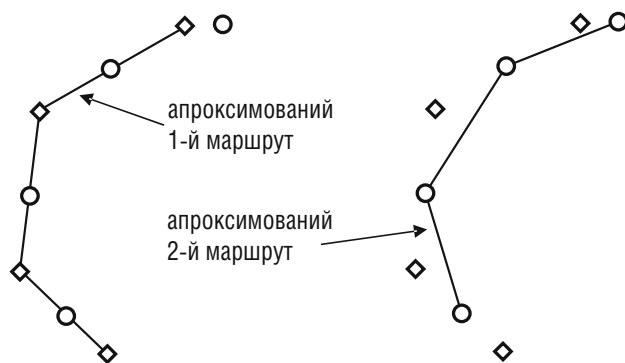


Рис. 1. Демонстрація некомутативності метрики (1)

Для забезпечення комутативності метрики остаточно відстань між маршрутами обчислюється таким чином

$$R_k(M_1, M_2) = \min(R(M_1, M_2), R(M_2, M_1)). \quad (3)$$

З використанням (3) для наявних даних про маршрути в базі геоданих ІТС «Гарт-12» можна побудувати матрицю відстаней R , елементи якої r_{ij} визначимо таким чином:

$$r_{ij} = \begin{cases} 0, & i = j \\ R^k(M_i, M_j), & i \neq j \end{cases} \quad (4)$$

З метою проведення класифікації наявних даних з визначенням типових маршрутів руху суден пропонується для кожного i -го рядка матриці R обчислити: кількість елементів у рядку n_s , які не перевищують порогового для кластеризації рівня r_{pk} (цей рівень отримується з використанням експертної оцінки); сумарну відстань до маршрутів s_v , відстані до яких не перевищують порогового для кластеризації рівня r_{pk} . Результати цих обчислень, які сортуються за зменшенням n_s (для елементів з однаковими значеннями n_s здійснюється сортування за збільшенням s_v), утворюють упорядковану множину M_s , яка містить елементи $(i, n_s(i), s_v(i))$.

Кластеризація маршрутів проводиться на основі M_s та R з використанням методу, який дає змогу, встановивши

належність даних про окремі маршрути до кластерів, визначити їх центри як маршрути, сумарна відстань від яких до інших маршрутів у кластері є мінімальною. Цей метод передбачає повторення таких дій до тих пір, поки потужність множини M_s є більшою за 0:

- якщо для першого елемента множини M_s $n_s(i) > 0$, то створюється новий кластер із центральним маршрутом i та до створеного кластеру додаються всі маршрути, відстань до яких від центрального маршруту не перевищує r_{pk} ;

- з множини M_s видаляється перший елемент i в разі виконання умови $n_s(i) > 0$ всі елементи, які відповідають маршрутам, доданим до кластера в попередньому кроці.

Метод кластеризації маршрутів, заданих в ІТС «Гарт-12» дискретними відліками, отримав програмно-алгоритмічну реалізацію (рис. 2).

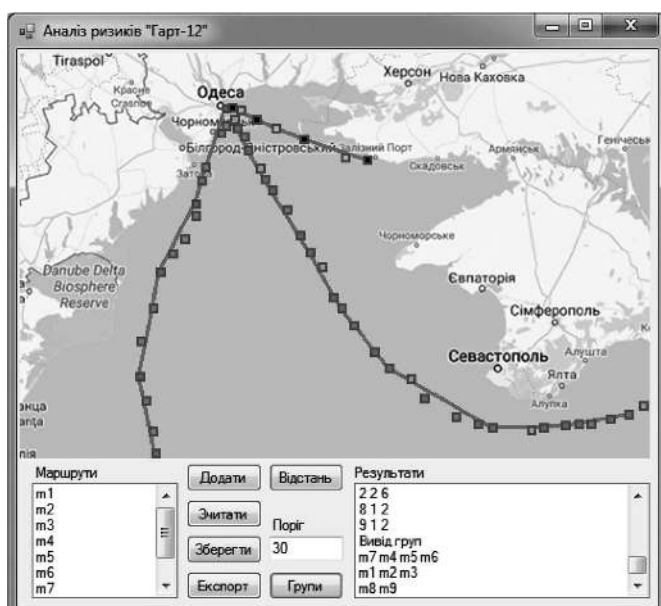


Рис. 2. Кластеризація даних про маршрути

На *рисунок 2* продемонстрований приклад класифікації даних про маршрути руху суден на основі кластеризації. У програмі передбачена можливість експорту отриманих типових маршрутів у форматі XML.

Наступною задачею, яку потрібно розв'язати для оцінювання ризиків порушення прикордонного законодавства на морській ділянці кордону, є встановлення просторової належності нових даних до одного з попередньо визначених типових маршрутів. Така належність визначається, виходячи з критерію неперевикнення відстані від досліджуваного до найближчого типового маршруту порогового рівня r_p . Цей пороговий рівень встановлюється на основі експертних оцінок.

При успішному встановленні просторової належності досліджуваного нового маршруту, інформація про який

надходить до СВНО, проводиться оцінювання ризиків за декількома факторами.

Першим фактором, який підлягає оцінюванню, є відповідність описових даних досліджуваного маршруту (порт відправки, порт призначення) еталонним, які відповідають типовому маршруту. Оскільки ці описові дані є текстовими, з метою зменшення впливу можливих незначних помилок оператора для їх нечіткого порівняння використовується текстова метрика Левенштейна [6]. Поріг можливих помилок, які не враховуються, визначений аналогічно до того, що використовується в програмно-технічному комплексі прикордонного контролю «Гарт-1/П». Позначимо першу ознаку логічною змінною o_1 .

Другим фактором, що впливає на прийняття рішення, є відстань від еталонного типового маршруту M_1 , належність до якого встановлена, до досліджуваного маршруту M_2 , обчислена за (1)–(2). Позначимо відношення цієї відстані до порогового рівня

$$o_2 = \frac{R(M_1, M_2)}{r_p}. \quad (5)$$

Слід зазначити, що ця відстань описує усереднене відхилення точок досліджуваного маршруту від апроксимованого типового. З метою оцінювання пікових відхилень введемо ще один параметр – максимальне відхилення, яке також оцінюватиметься у відношенні до порогового рівня

$$o_3 = \frac{\max(\min(R_{minij}))}{r_p}. \quad (6)$$

На основі дослідження роботи експертів щодо виявлення ризиків порушення прикордонного законодавства на основі аналізу даних у системі висвітлення надводної обстановки були визначені три категорії ризику (низький, середній і високий) та отримані критерії для їх встановлення. Низькому рівню ризиків відповідають випадки, коли описи маршрутів збігаються, показник $o_2 < 0.5$ і показник $o_3 < 0.8$. Високий рівень ризику настає за умови нерівності опису маршрутів, або, за умови їх рівності, при $o_2 > 0.75$ або $o_3 > 0.9$. У всіх інших випадках рівень ризику середній.

На основі наведеної вище інформації сформовані логічні вирази для встановлення ризиків порушення прикордонного законодавства, які наведені в *таблиці 1*.

Таблиця 1

Логічні вирази для встановлення ризиків

Ризик	Логічний вираз
Низький	$o_1 \cap (o_2 < 0.5) \cap (o_3 < 0.8)$
Середній	$o_1 \cap \overline{o_2 > 0.75} \cap \overline{o_3 > 0.9} \cap (\overline{o_2 < 0.5} \cup \overline{o_3 < 0.8})$
Високий	$\overline{o_1} \cup (o_1 \cap ((o_2 > 0.75) \cup (o_3 > 0.9)))$

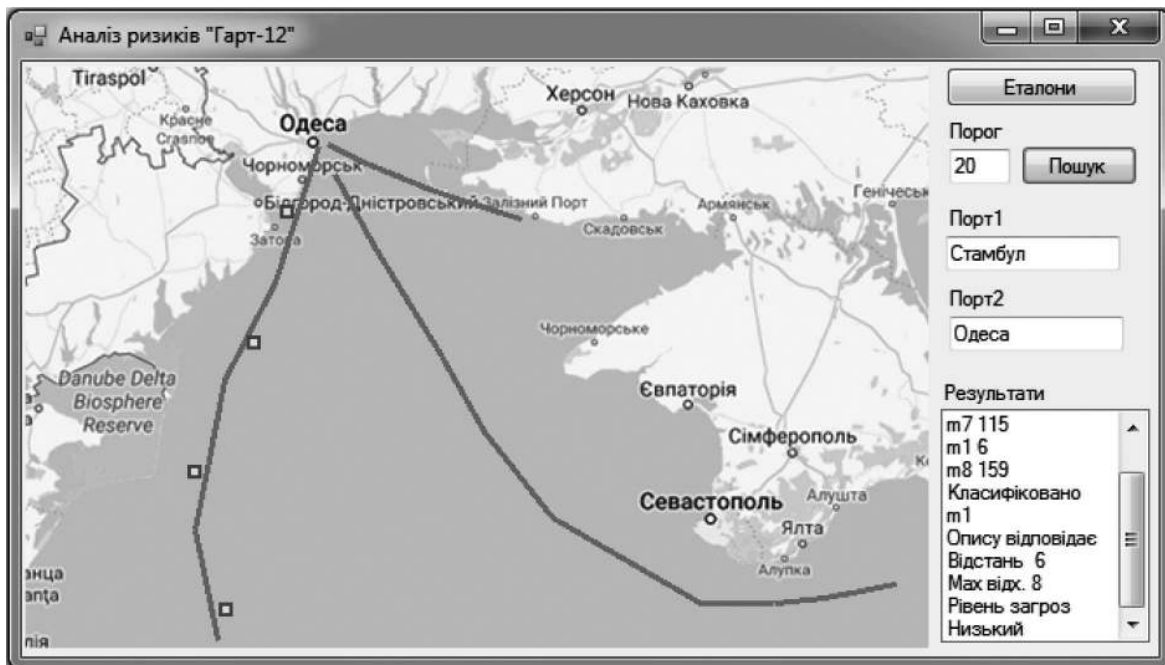


Рис. 3. Програмне забезпечення встановлення ризиків на основі просторового аналізу даних про маршрути руху

Описана вище методика визначення показників оцінювання ризиків і встановлення ризиків на основі логічних виразів була реалізована в програмному забезпеченні, робота якого представлена на *рисунку 3*.

Як видно з *рисунка 3*, для досліджуваного варіанта за внесеними даними маршрут було класифіковано й віднесено до типового маршруту « m_1 ». Результати нечіткого порівняння текстових описових даних введеного маршруту (порти відправки та призначення) і даних визначеного типового маршруту не перевищили припустимого порогу помилки. Відстань між маршрутами склала 6 км, що відповідає $o_2 = 0.3$, тобто $o_2 < 0.5$. Максимальне відхилення досліджуваних даних від типового маршруту склало 8 км. Такому максимальному відхиленню відповідає $o_3 = 0.4$, тобто $o_3 < 0.8$. Тому у програмному забезпеченні, відповідно до логічних виразів у *таблиці 1*, визначений рівень загроз «низький».

При суттєвому відхиленні одного з відліків або більшому середньому віддаленні точок досліджуваного маршруту рівень ризиків зростає (*рис. 4*).

Висновки. У роботі розроблена методика виявлення просторових аномалій руху суден та її використання при оцінюванні ризиків в ІТС морської охорони «Гарт-12». Важливим елементом цієї методики є метод кластеризації маршрутів, заданих в ІТС «Гарт-12» дискретними відліками, який базується на використанні метрики (3). На основі кластеризації запропоноване проведення класифікації наявних даних у базі ІТС «Гарт-12» та визначення типових маршрутів. Дані про типові маршрути дають підстави класифікувати нові дані, які надходять до

системи висвітлення надводної обстановки, та проводити на їх основі аналіз ризиків.

У роботі визначені логічні вирази для встановлення рівня загроз порушення прикордонного законодавства на морській ділянці кордону.

Напрямом подальших досліджень є формування підходів до виявлення часових аномалій руху морських суден.

Перелік літератури

1. Боровик О. В. Сучасні підходи до геомодельовання з використанням ARCGIS / О. В. Боровик, І. І. Балицький, Р. В. Рачок // Збірник наукових праць Національної академії Державної прикордонної служби України. – Серія : Військові та технічні науки / голов. ред. Б. М. Олексієнко. – Хмельницький : Вид-во НАДПСУ, 2015. – № 4 (66). – С. 275–282.
2. Рачок Р. В. Формування концепції розбудови геоінформаційної системи у сфері забезпечення прикордонної безпеки / Р. В. Рачок // Збірник наукових праць Національної академії Державної прикордонної служби України. – Серія : Військові та технічні науки / голов. ред. Б. М. Олексієнко. – Хмельницький : Вид-во НАДПСУ, 2017. – № 2 (72). – С. 295–310.
3. Боровик О. В. Визначення підходів до просторового аналізу даних в інформаційно-телекомунікаційній системі морської охорони «Гарт-12» / О. В. Боровик, Р. В. Рачок, В. Ю. Мазур // Збірник наукових праць Національної академії Державної прикордонної служби України. – Серія : Військові та технічні науки / голов. ред. Б. М. Олексієнко. – Хмельницький : Вид-во НАДПСУ, 2018. – № 1 (75). – С. 295–310.
4. Олдендерфер М. С., Блешфилд Р. К. Кластерный анализ // Факторный, дискриминантный и кластерный анализ / М. С. Олдендерфер, Р. К. Блешфилд; пер. с англ. под. ред. И. С. Енюкова. – М. : Финансы и статистика, 1989. – 215 с.

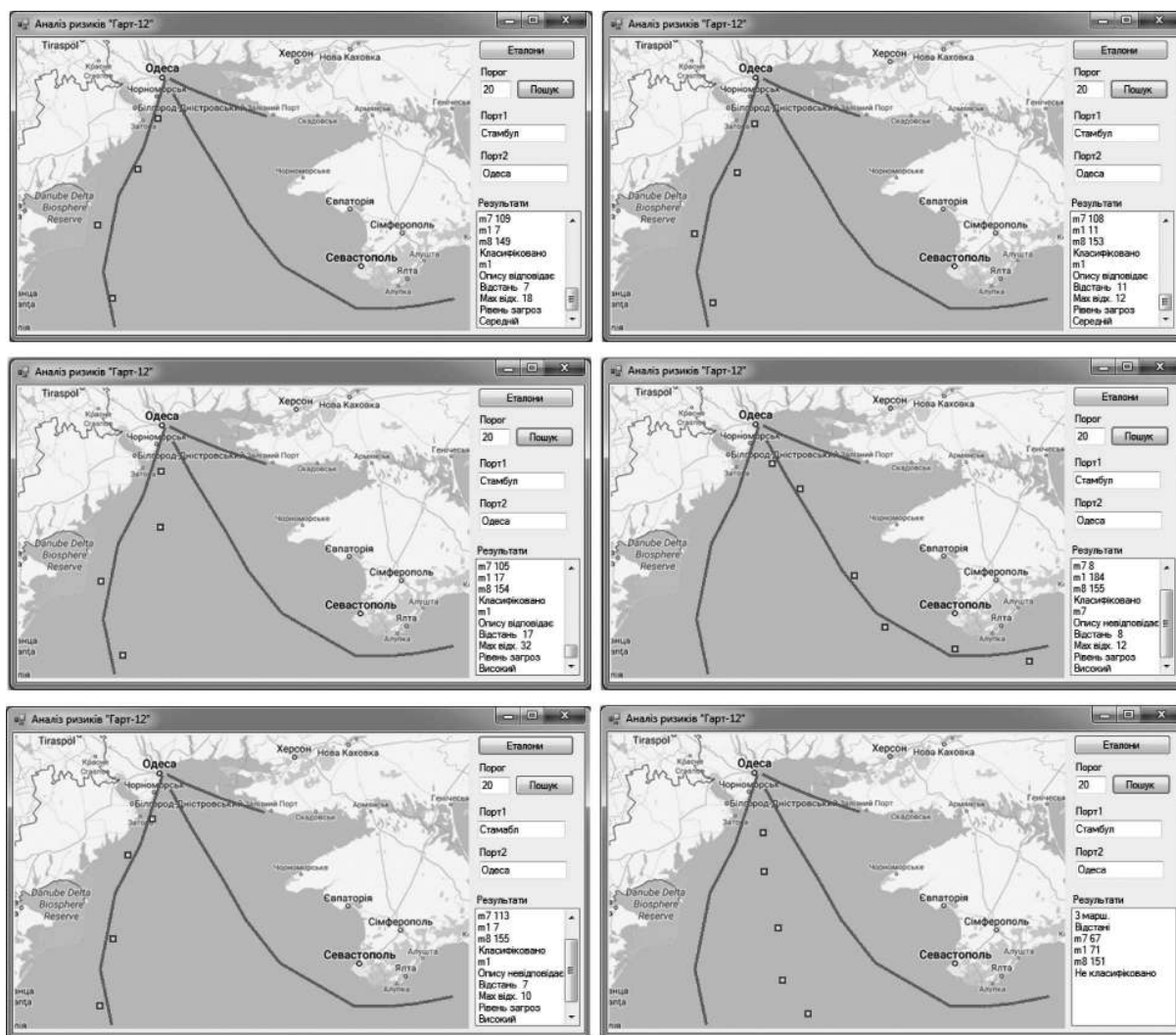


Рис. 4. Різні варіанти збільшення ризиків

5. Боровик О. В., Лемешко В. В. SWOT-анализ как информационная технология планирования национальной безопасности на государственной границе / О. В. Боровик, В. В. Лемешко // Материалы Международной научно-практической конференции (29–30 октября 2012 года) «Информационные и коммуника-

ционные технологии в науке, образовании и производстве». – Тирасполь : Ликрис, 2012. – С. 66–69.

6. Левенштейн В. И. Двоичные коды с исправлением выпадений, вставок и замещений символов / В. И. Левенштейн // ДАН СССР. – 1965. – Т. 163. – № 4. – С. 845–848.