

УДК 528.29

Валентин МАЗУР,
кандидат військових наук, доцент,
Національна академія Державної прикордонної служби України
імені Богдана Хмельницького, м. Хмельницький

Олег БОРОВИК,
доктор технічних наук, професор,
Національна академія Державної прикордонної служби України
імені Богдана Хмельницького, м. Хмельницький

Роман РАЧОК,
кандидат технічних наук, доцент,
Національна академія Державної прикордонної служби України
імені Богдана Хмельницького, м. Хмельницький

ВИЗНАЧЕННЯ ПІДХОДІВ ДО ПРОСТОРОВОГО АНАЛІЗУ ДАНИХ В ІНФОРМАЦІЙНО- ТЕЛЕКОМУНІКАЦІЙНІЙ СИСТЕМІ МОРСЬКОЇ ОХОРОНИ "ГАРТ-12"

У дослідженні сформульована задача аналізу даних в інформаційно-телекомунікаційній системі Морської охорони "Гарт-12". Потреба у просторовій класифікації даних про рух суден обумовила необхідність розробки маршрутної метрики. У випадку лінійної апроксимації маршруту отримані вирази для обчислення маршрутної метрики.

Ключові слова: *система висвітлення надводної обстановки, інформаційно-телекомунікаційна система Морської охорони “Гарт-12”, метрика, маршрут, критерій.*

Постановка проблеми у загальному вигляді. Важливим чинником сталого розвитку нашої суверенної держави є надійна охорона кордону. У сучасних умовах постійного зростання обсягів даних, які необхідно врахувати при прийнятті рішень, надзвичайно важливою є інформаційна складова сучасної моделі охорони кордону. Ця складова реалізується з використанням інтегрованої інформаційно-телекомунікаційної системи (ІТС) “Гарт”. Однією з важливих складових “Гарт” є інформаційно-телекомунікаційна система (ІТС) Морської охорони “Гарт-12”.

Однак на сьогодні “Гарт-12” використовується лише для висвітлення надводної обстановки в Азово-Чорноморській акваторії. При цьому існує можливість отримання інформації з ІТС “Гарт-12” про поточне положення суден, порт відправки та порт призначення, тип вантажу та інші додаткові дані. При виборі судна у системі висвітлення надводної обстановки існує можливість перегляду попередніх точок його маршруту, отриманих у певні дискретні моменти часу. При цьому візуалізація маршруту здійснюється з використанням лінійної апроксимації (наявні точки з’єднуються відрізками прямих).

Слід зазначити, що по такій, наявній в межах системи висвітлення надводної обстановки інформації, навіть експертам може бути важко визначити потенційні загрози.

Отже, наявність значних обсягів даних щодо надводної обстановки у ІТС “Гарт-12” та складність аналізу ризиків на їх основі робить важливою автоматизацію їх обробки.

Аналіз останніх досліджень і публікацій, в яких започатковано вирішення даної проблеми та на які опираються автори. Основою ІТС “Гарт-12” є геоінформаційна система Державної прикордонної служби України “Гарт-17”. У ряді робіт [1–4] описані задачі прикордонного відомства, що потребують геообробки даних. Однією з цих задач є виявлення аномалій руху морських та річкових суден і на їх основі

прогнозування небезпек в ІТС “Морська охорона” (“Гарт-12”). Однак підходів до її вирішення у цих дослідженнях не наводиться.

З метою проведення аналізу даних про маршрути морських та річкових суден можливо використати методи класифікації, які описані зокрема у [5–6]. Однак особливості представлення даних про ці маршрути у ІТС “Гарт-12” потребують розвитку базового науково-методичного апарату класифікації. Незбігання у часі дискретних відліків про положення суден вимагають визначення підходів до порівняння маршрутів, що є основою для їх подальшої класифікації і знаходження аномалій.

Уникнути цих складнощів можливо з використанням методів машинного навчання, наприклад, штучних нейронних мереж [7]. Проте ці методи є чутливими до багатьох факторів: способу кодування даних, архітектури нейромережі, її навчання та ін. В окремих випадках їх використання може приводити до отримання хибних рішень.

Використання методів дискретного аналізу, кластеризації потребує встановлення міри подібності об’єктів, які класифікуються. Усе це обумовлює потребу визначення метрики для маршрутів суден з урахуванням особливості їх представлення у ІТС “Гарт-12”.

Мета статті – визначення підходів до проведення аналізу даних в ІТС “Гарт-12”.

Виклад основного матеріалу дослідження. Як було зазначено у [2–3], основою аналізу ризиків у “Гарт-12” може стати виявлення аномалій руху морських та річкових суден.

В аналізі даних під виявленням аномалій мається на увазі знаходження та ідентифікація елементів, подій або спостережень, що не відповідають очікуваній поведінці або іншим елементам набору даних [8]. У цьому контексті, при аналізі даних у ІТС Морської охорони необхідно вирішити низку часткових завдань: формування множини типових маршрутів на основі аналізу даних у базі даних “Гарт-12”; проведення класифікації нових маршрутів; виявлення аномалій та оцінка ризиків на їх основі.

Для кожного маршруту в базі даних “Гарт-12” міститься інформація про порт відправлення, порт призначення та тип вантажу. Ці

дані можливо використовувати для попередньої класифікації маршруту. Тоді після подальшої просторової класифікації тотожність цих результатів може бути використана при оцінці ризиків. Проте з часом типові маршрути можуть змінюватись за рахунок появи нових проміжних місць зупинок. Тому одному маршруту за попередньою класифікацією можуть відповідати декілька просторових маршрутів. Належність досліджуваного маршруту до одного з них не є підставою для зростання ризику.

При пересуванні кожного судна в базі даних “Гарт-12” в окремі дискретні моменти часу фіксуються його координати. Позначимо сукупність цих координат для i -го судна впорядкованою множиною (кортежем) $M_i = \{(x_1, y_1), \dots, (x_{k_i}, y_{k_i})\}$, де k_i – потужність множини M_i .

Визначення аномалій руху та проведення класифікації маршрутів потребує встановлення ступеня їх подібності, який можливо визначити на основі введення відповідної метрики. Відстань між двома маршрутами M_1 та M_2 позначимо $R(M_1, M_2)$.

Цю відстань можливо було б визначити на основі стандартної метрики між окремими точками множин M_i , що їх представляють. Однак відмінності в потужності цих множин ускладнюють таке визначення метрики. Основною проблемою цього підходу є те, що окремі точки представлення маршрутів отримані в різний час. Навіть у випадку, коли два маршрути збігаються, відстані між окремими дискретними точками, що їх описують, можуть бути суттєвими (рис. 1).

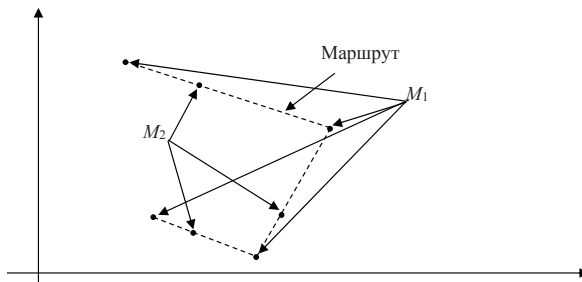


Рис. 1. Відмінності при представленні маршруту, обумовлені різницею відліків у часі

Відповідно до рис. 1 перший маршрут представлений множиною, що містить чотири елементи. Другий маршрут на рис. 1 описується множиною з трьох елементів. Усі точки цих маршрутів лежать на одній пунктирній лінії. Оскільки ці відліки робились в різні моменти часу, відстані між ними є суттєвими. У зв'язку з цим будувати метрику на основі відстаней між окремими дискретними відліками є некоректним.

Ураховуючи вказане вище, для пошуку відстані між двома маршрутами, представленими дискретними впорядкованими множинами відліків, пропонується один з цих маршрутів відновити з використанням апроксимації між кожною парою сусідніх наявних точок. Нехай: M_1 – множина, яка описує відновлюваний маршрут, $f_k(x)$ – функція, яка апроксимує перший маршрут між точками (x_k, y_k) і (x_{k+1}, y_{k+1}) . Тоді, з метою визначення метрики з точки зору геометричної близькості маршрутів, для будь якого відліку, що описує другий маршрут, пропонується знайти найближчу точку апроксимованого першого маршруту (рис. 2).

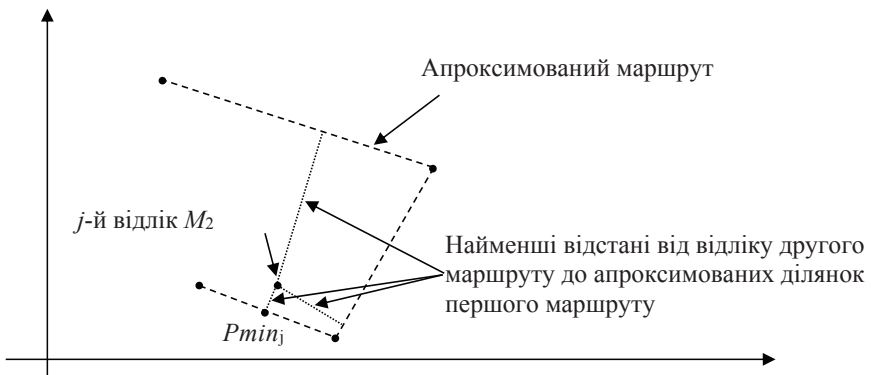


Рис. 2. Використання апроксимації маршруту для визначення метрики

Для простоти на рис. 2 подана апроксимація з використанням лінійної залежності. Однак в загальному випадку можливо обрати і більш складну функцію. З прикладу (рис. 2), досліджуваній j -й точ-

ці другого маршруту відповідатиме точка першого маршруту $Pmin_j$, мінімальна відстань від якої до трьох апроксимованих ділянок виявилась найменшою. Цю відстань позначимо $Rmin_j$. Такі відстані і пропонується взяти за основу при визначенні метрики $R(M_1, M_2)$, яку визначимо так

$$R(M_1, M_2) = \frac{\sum_{j=1}^{k_2} Rmin_j}{k_2}. \quad (1)$$

Звичайно, основним проблемним питанням при обчисленні (1) є розрахунок величин $Rmin_j$, для якого необхідне визначення точок $Pmin_j$. Розглянемо можливий підхід до вирішення цієї задачі.

Визначимо для i -го апроксимованого сегменту першого маршруту точку (x, y) , для якої забезпечується найменша відстань до одного з дискретних відліків другого маршруту (x_i, y_i) .

Відстань до довільної точки досліджуваного сегменту обчислюється за виразом

$$R(x) = \sqrt{(x - x_1)^2 + (y_1 - f_i(x))^2}. \quad (2)$$

Тому для пошуку мінімального значення необхідно прирівняти до нуля похідну підкореневого виразу і розв'язати рівняння

$$\frac{1}{2} \cdot \frac{d}{dx} f_i^2(x) - y_1 \cdot \frac{d}{dx} f_i(x) + x - x_1 = 0. \quad (3)$$

У випадку, якщо рівняння (3) матиме декілька розв'язків, необхідно визначити остаточне рішення за критерієм мінімальності (2).

При використанні лінійної апроксимації

$$f_i(x) = k_i \cdot x + b_i. \quad (4)$$

За цих умов (2) матиме вигляд функції з одним мінімумом (рис. 3).

При цьому розв'язок (3) має бути єдиним і відповідати шуканому мінімуму. Це дозволяє не здійснювати перевірки на мінімальність отриманого екстремуму.

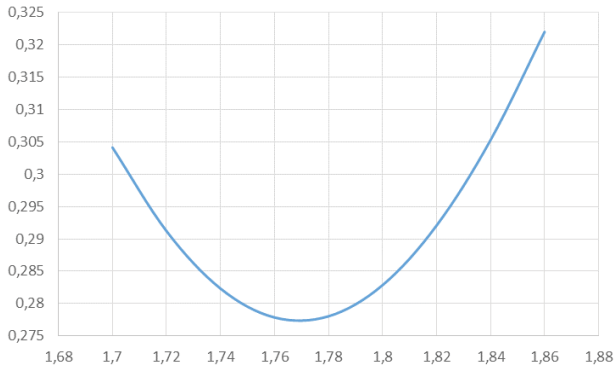


Рис. 3. Залежність відстані при лінійній апроксимації

Підставивши (4) у (3) і розв'язавши отримане рівняння, отримуємо

$$x = \frac{y_1 \cdot k_i + x_1 - b_i \cdot k_i}{1 + k_i^2}. \quad (5)$$

З урахуванням (4) можливо отримати шукану точку (x, y) .

Слід відмітити, що у випадку лінійної апроксимації вираз (5) можна отримати, виходячи з більш простих міркувань. Очевидно, що шукана точка відповідатиме точці перетину (4) з перпендикулярною до неї прямою, яка проходить через точку (x_1, y_1) . Ця пряма описується рівнянням

$$y - y_1 = -\frac{1}{k_i}(x - x_1). \quad (6)$$

Точка перетину (4) і (6) визначається з системи

$$\begin{cases} y = k_i \cdot x + b_i \\ y - y_1 = -\frac{1}{k_i}(x - x_1) \end{cases} \quad (7)$$

Розв'язавши (7), також можна отримати (5).

Отже, вирази (4)-(5) та (2) дають можливість визначити мінімальні відстані від кожної j -ї точки другого маршруту до від-

різків, якими апроксимується перший маршрут. Найменша з цих відстаней R_{min_j} використовується у виразі (1) для обчислення остаточної метрики встановлення міри подібності між маршрутами.

Висновки. У роботі розглянуті можливі підходи до просторового аналізу даних про маршрути суден в ІТС Морської охорони “Гарт-12”. Показано доцільність визначення особливої метрики для класифікації маршрутів, які представлені в ІТС “Гарт-12” впорядкованими множинами точок, отриманих у довільні моменти часу. У випадку лінійної апроксимації даних отриманий вираз для обчислення відстані, яка відображає ступінь подібності маршрутів. Отримані вирази можна використовувати для класифікації маршрутів суден у ІТС “Гарт-12”, визначення геометричних відхилень маршрутів суден з метою проведення подальшої оцінки ризиків.

Напрямами подальших досліджень є формалізація та алгоритмізація методики визначення відстаней між маршрутами суден в “Гарт-12”, встановлення критеріїв подібності цих маршрутів та розробка методики оцінки ризиків у системі Морської охорони кордону.

Список використаної літератури

1. Боровик О. В., Балицький І. І., Рачок Р. В. Сучасні підходи до геоделювання з використанням ARCGIS. *Збірник наукових праць Національної академії Державної прикордонної служби України. Серія : військові та технічні науки /* голов. ред. Олексієнко Б. М. Хмельницький : Вид-во НАДПСУ, 2015. № 4(66). С. 275–282.
2. Рачок Р. В. Формування концепції розбудови геоінформаційної системи у сфері забезпечення прикордонної безпеки. *Збірник наукових праць Національної академії Державної прикордонної служби України. Серія : військові та технічні науки /* голов. ред. Олексієнко Б. М. Хмельницький : Вид-во НАДПСУ, 2017. № 2(72). С. 295–310.
3. Рачок Р. В. Проблемні питання розбудови геоінформаційної системи Державної прикордонної служби України. *Освітньо-наукове забезпечення діяльності правоохоронних органів і військових формувань України : тези VIII Всеукраїнської науково-практичної конференції (Хмельницький, 10 грудня 2015 р.)*. Хмельницький, 2015. С. 519.

4. Боровик О. В., Рачок Р. В. Сучасні можливості геоделювання та напрями розвитку геоінформаційної системи Державної прикордонної служби України на основі використання ARCGIS. *Перспективи розвитку озброєння та військової техніки Сухопутних військ* : Міжнар. наук.-техн. конф. (Львів, 18–20 травня 2016 р.). Львів, 2016. С. 171.

5. Шлезингер М., Главач В. Десять лекцій по статистическому и структурному распознаванию. Киев: Наукова думка, 2004.

6. Журавлев Ю. И., Рязанов В. В., Сенько О. В. “Распознавание”. Математические методы. Программная система. Практические применения. Москва : Фазис, 2006.

7. Kruse Rudolf, Borgelt Christian, Klawonn F., Moewes Christian, Steinbrecher Matthias, Held Pascal (2013). Computational intelligence : a methodological introduction. Springer. ISBN 9781447150121. OCLC 837524179.

8. Chandola V., Banerjee A., Kumar, V. (2009). Anomaly detection: A survey. *ACM Computing Surveys*[en]} 41 (3). С. 1. doi:10.1145/1541880.1541882

Мазур В., Боровик О., Рачок Р. Определение подходов к пространственному анализу данных в информационно-телекоммуникационной системе морской охраны “Гарт-12”

В исследовании сформулирована задача анализа данных в информационно-телекоммуникационной системе Морской охраны “Гарт-12”. Потребность в пространственной классификации данных о движении судов обусловила необходимость разработки маршрутной метрики. В случае линейной аппроксимации маршрута получены определения для вычисления маршрутной метрики.

Ключевые слова: система освещения надводной обстановки, информационно-телекоммуникационная система Морской охраны “Гарт-12”, метрика, маршрут, критерий.

Mazur V., Borovik O., Rachok R. Determination of approaches to spatial analysis of data in the information and telecommunication system of Marine guard “Gart-12”.

In the study task of analysis is formulated in informational-telecommunication systems of Marine guard “Gart-12”. The need for a spatial classification of vessel traffic data necessitated the development of

a route metric. In the case of the route, the expressions for calculating the routing metric are obtained.

Keywords: *system of illumination of surface environment, informational-telecommunication systems of Marine guard "Gart-12", metrics, route, criterion.*