
ВІЙСЬКОВІ НАУКИ

УДК 37.04:004 (355)

Володимир АНДРУШКО,
кандидат військових наук,
Національна академія Державної прикордонної служби України
імені Богдана Хмельницького, м. Хмельницький

ПОРЯДОК ЗАСТОСУВАННЯ ПАРАМЕТРІВ ШТУЧНОЇ НЕЙРО-НЕЧІТКОЇ МОДЕЛІ ПРОГНОЗУВАННЯ ІНТЕНСИВНОСТІ РУХУ ОСІБ ТА ТРАНСПОРТНИХ ЗАСОБІВ

У статті представлено рекомендації начальнику відділу прикордонної служби з прийняття рішень на основі прогнозування із застосуванням моделі штучної нейро-нечіткої мережі щодо інтенсивності діяльності пунктів пропуску Державної прикордонної служби України. Основні з них стосуються роботи з інформацією (даними) і планування зміни прикордонних нарядів з урахуванням результатів прогнозу. Дослідження показали, що даний підхід дає можливість обґрунтовувати кількісний склад змін прикордонних нарядів у пунктах пропуску.

Ключові слова: прогноз, рекомендації, нейронна мережа, планування.

Постановка проблеми у загальному вигляді. Однією з найважливіших частин розробки управлінського рішення є прогнозування.

Поява у Державній прикордонній службі України (ДПСУ) потужних засобів збору й обробки інформації у складі інформаційно-телекомунікаційних систем “Гарт”, які містять у базах даних відомості стосовно осіб, автотранспортних засобів, що перетнули державний кордон (ДК), випадків порушення законодавства з питань перетинання ДК та інших показників, надає можливість автоматизувати процес прогнозування. Більшість цих даних можна інтерпретувати як часовий (чисельний) ряд.

Побудова адекватної моделі прогнозу передбачає вибір параметрів прогнозування та попередню обробку їх значень.

Маючи оброблені вихідні дані, далеко не завжди відразу можна побудувати адекватну модель прогнозу, оскільки не всі параметри, за допомогою яких охарактеризована інтенсивність руху через ДК, на Кроці побудови апріорної концептуальної моделі мають значення для побудови моделі, крім того, деякі параметри можуть чинити негативний вплив на модель.

Крім фільтрації і згладжування даних, на практиці часто необхідно знайти значення деяких функцій, що характеризують певні властивості даних (наприклад, математичне сподівання, дисперсію), які будуть необхідні на наступних Кроках прогнозу, тобто необхідно виконати статистичну обробку експериментальних даних.

На наступному Кроці необхідно оцінити інформативність (значимість) параметрів (ознак) прогнозу і відібрати найбільш значимі з нього для використання на наступних Кроках прогнозу. Виконання цього Кроку дозволяє не тільки поліпшити модель, виключивши шкідливі і непотрібні параметри, але спростити і здешевити процес прогнозу за рахунок скорочення вимірювань параметрів.

На жаль, при побудові концептуальної моделі прогнозу не завжди вдається вибрати такий набір ознак, який був би тісно пов’язаний з тими параметрами моделі, моделювання яких необхідно здійснити. Найчастіше деякі параметри, які тісно пов’язані з модельованими параметрами об’єкта пізнання, у набір ознак безпосередньо не включаються, а присутні в ньому побічно через інші параметри пов’язані з

ними. Тому для з'ясування наявності таких прихованих параметрів і визначення їх зв'язку з параметрами, що входять у концептуальну модель, необхідно здійснювати Крок факторного аналізу.

Аналіз останніх досліджень і публікацій, в яких започатковано вирішення даної проблеми та на які опирається автор. Загальне завдання прогнозування показників службової діяльності прикордонних структур частково розглядалось А. Ф. Кучковим та Н. Ф. Лукашевичем [1]. Завданням прогнозування на основі моделей часових (чисельних) рядів присвячено низку досліджень із використанням статистичних методів і суб'єктивних знань експертів, зокрема роботи Дж. Бокса, Г. Дженкінса, В. П. Боровікова, Г. І. Івченко та ін. [2; 3]. Недоліками цих підходів є: відсутність у моделей відомостей щодо структури й системи зв'язків реального об'єкта; труднощі побудови моделей за умови, що дані зберігаються в різних часових рядах і мають часові зміщення щодо один одного; недостатня точність прогнозу; значна чутливість отриманих результатів до недостатньої інформації; залежність результату прогнозу від кваліфікації аналітика в конкретній предметній сфері. Зазначені недоліки можна долати із застосуванням апарата штучних нейрон-нечітких мереж (ШННМ), дослідження яких здійснено у [4].

У той же час відсутні дослідження стосовно вибору параметрів прогнозування із застосуванням ШННМ в оперативно-службовій діяльності.

Метою статті є розробка методичного підходу щодо вибору параметрів прогнозування інтенсивності руху осіб і транспортних засобів у пунктах пропуску ДПСУ із застосуванням ШННМ.

Виклад основного матеріалу дослідження. Розроблено та запропоновано формалізований підхід до вибору параметрів для прогнозування. Методологія вибору сукупності параметрів заснована на використанні факторного аналізу.

Об'єктом дослідження є інтенсивність руху осіб/транспортних засобів, стан якої можна описати сукупністю параметрів, що являють

собою вектор “службового” стану $\bar{Y} = (y_1, y_2, \dots, y_n)$, де n – число параметрів системи.

Наприклад, таких як кількість осіб, які перетнули пункт пропуску за одну добу, день тижня (включаючи святкові дні), різниця в ціні на “контрабандні товари” (сигарети, горілка, цукор, м’ясо тощо), курс валют, наявність у суміжній країні надзвичайної ситуації тощо.

Основним завданням є перехід від опису досліджуваної системи, заданої великим набором вимірюваних параметрів до опису меншим числом максимально інформативних ознак, що відображають найбільш істотні властивості системи. Передбачається, що кожен з параметрів y_i , що описує стан системи, може бути представлений як функція невеликого числа загальних факторів p_i і характерного фактора u_i :

$$y_i = f(p_1, p_2, \dots, p_r, u_i). \quad (1)$$

Загальні фактори являють собою розрахункові змінні, тобто нові характеристики системи. Вихідною інформацією для проведення факторного аналізу є кореляційна матриця параметрів. У подальшому використані такі умовні позначення: j – номер об’єкта пропуску (особи, транспортні засоби, вантажі) ($j = 1, \dots, m$); i – номер параметра контролю ($i = 1, \dots, n$); k – номер виміру параметрів ($k = 1, \dots, s$).

Крок 1. Побудова кореляційної матриці $R_{ij} = \|r_{ik}\|$. На основі коваріаційної матриці будується кореляційна матриця параметрів $R_{ij} = \|r_{ik}\|$. Коефіцієнти парних кореляцій змінних розраховуються за такою формулою:

$$R_{ij} = \frac{C_{ij}^*}{\sqrt{C_{ii}^* \cdot C_{jj}^*}}, \quad (2),$$

де $C_{ij}^*, C_{jj}^*, C_{ii}^*$ – коефіцієнти коваріації.

Отримана кореляційна матриця параметрів $R_{ij} = \|r_{ik}^h\|$ має такі значення елементів:

$$R_{ij} = \begin{pmatrix} 1 & 0,920408 & 0,88388 & -0,9264 & -0,8113 & 0,96709 \\ 0,920408 & 1 & 0,925029 & -0,98724 & -0,66375 & 0,917863 \\ 0,88388 & 0,925029 & 1 & -0,97285 & -0,72535 & 0,951378 \\ -0,9264 & -0,98724 & -0,97285 & 1 & 0,718748 & -0,95459 \\ -0,8113 & -0,66375 & -0,72535 & 0,718748 & 1 & -0,87373 \\ 0,96709 & 0,917863 & 0,951378 & -0,95459 & -0,87373 & 1 \end{pmatrix}$$

Крок 2. Побудова редуційованої кореляційної матриці $R_h = \|r_{ik}^h\|$, де $i, k = 1, \dots, n$. Для цього попередні оцінки спільностей h_i^2 отримують на основі обчислення квадрата множинної кореляції для кожної змінної за допомогою оберненої матриці R^{-1} за такою формулою:

$$h_i^2 = R_i^2 = 1 - \frac{1}{r^{ii}}, \quad (3)$$

де r^{ii} – діагональний елемент зворотної матриці.

Редуційована кореляційна матриця $R_h = \|r_{ik}^h\|$ набуде вигляду:

$$R_h = \begin{pmatrix} 0,999893 & 0,920408 & 0,883876 & -0,926420 & -0,811290 & 0,967092 \\ 0,920408 & 0,999997 & 0,925029 & -0,987240 & -0,663750 & 0,917863 \\ 0,883876 & 0,925029 & 0,999989 & -0,972850 & -0,725350 & 0,951378 \\ -0,926420 & -0,987240 & -0,972850 & 0,999999 & 0,718748 & -0,954590 \\ -0,811290 & -0,663750 & -0,725350 & 0,718748 & 0,989588 & -0,873730 \\ 0,967092 & 0,917863 & 0,951378 & -0,954590 & -0,873730 & 0,999972 \end{pmatrix}$$

Крок 3. Процедура виділення факторів і перехід до матриці факторного відображення $A = \|a_{il}\|$, елементами якої є факторні навантаження, де $i = 1, \dots, n$ – параметри; $l = 1, \dots, r$ – фактори.

Кореляційна матриця може бути відтворена за допомогою факторного відображення і кореляцій між факторами:

$$\begin{pmatrix} h_1^2 & r_{12} & \dots & r_{1n} \\ r_{21} & h_2^2 & \dots & r_{2n} \\ \dots & \dots & \dots & \dots \\ r_{n1} & \dots & \dots & h_n^2 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1r} \\ a_{21} & a_{22} & \dots & a_{2r} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nr} \end{pmatrix} \times \begin{pmatrix} a_{11} & a_{21} & \dots & a_{n1} \\ a_{12} & a_{22} & \dots & a_{n2} \\ \dots & \dots & \dots & \dots \\ a_{1r} & a_{2r} & \dots & a_{nr} \end{pmatrix} \quad (4)$$

або у матричному вигляді:

$$R_h = A \cdot A',$$

де R_h – редуційована кореляційна матриця; h_i^2 – значення спільності; A – матриця факторного відображення; A' – транспонована матриця факторного відображення, a_{il} – факторні навантаження.

Отримання результату рівняння (4) полягає у визначенні власних значень λ_i і власних векторів α_{il} матриці кореляцій.

Для знаходження векторів α_{il} і значень λ_i редуційованої кореляційної матриці $R_h = \|r_{ik}^h\|$ використовується ітераційний метод. Ітераційний процес починається з вибору вектора $\alpha(1)$, елементи якого є першими наближеннями значень елементів власних векторів:

$$\beta^{(1)} = R_h \alpha^{(1)}, \quad (5)$$

$$\alpha^{(2)} = \beta^{(1)} / \max(\beta_i^{(1)}), \quad (6)$$

$$\alpha^{(k+1)} = \beta^{(k)} / \max(\beta_i^{(k)}), \quad (7)$$

де верхні індекси в дужках позначають крок ітерації.

Формула (7) є загальною для k кроків ітерації.

Процес повторюється до тих пір, поки не досягаються збіжності до першого власного значення матриці $R_h = \|r_{ik}^h\|$: $\lambda_1 = \max(\beta_i^{(k)})$ і відповідного першого власного вектора $\alpha^{(k)}$. Ітераційний процес закінчується, коли $\alpha^{(k)}$ і $\alpha^{(k-1)}$ із достатньою точністю збігаються один з одним, тобто $|\alpha^{(k)} - \alpha^{(k-1)}| \leq \varepsilon$. Для власного вектора $\alpha_{i1} = (\alpha_{11}, \alpha_{21}, \dots, \alpha_{n1})$ знайдені такі значення елементів: $\alpha_{i1} = (0,973435613; 0,959721698; 0,966355094; -0,98469688; -0,838327063; 1)$, а також перше власне значення $\lambda_1 = \max(\beta^{(4)})$, $\lambda_1 = 5,41408866$.

Крок 4. Обчислення навантажень першого фактора a_{il} матриці факторного відображення $A = \|a_{il}\|$. Навантаження матриці факторного

відображення $A = \|a_{il}\|$ можуть бути визначені за компонентами власних векторів матриці $R_h = \|r_{ik}^h\|$ за формулою:

$$a_{il} = \alpha_{il} \sqrt{\lambda_l} / \sqrt{\alpha_{i1}^2 + \alpha_{i2}^2 + \dots + \alpha_{in}^2}, \quad (8)$$

де $i, l = 1, 2, \dots, n$.

Отримані навантаження першого фактора: $a_{i1} = (0,968013955; 0,954376421; 0,960972872; -0,979212501; -0,833657907; 0,994430389)$.

Крок 5. Визначення матриці відтворених кореляцій R^+ з урахуванням тільки першого фактора і залишкової матриці кореляцій R_l .

Матриця відтворених кореляцій R^+ визначається за формулою:

$$R^+ = a_1 \cdot a_1', \quad (9)$$

де a_1 – вектор стовпчик навантажень першого фактора, a_1' – вектор стрічка навантажень першого фактора.

Остаточну матрицю кореляцій R_l , яка залишається після виділення першого фактора, визначаємо за такою формулою:

$$R_l = R_h - R^+. \quad (10)$$

Крок 6. Виділення другого чинника і визначення другого власного значення λ_2 і другого власного вектора α_{i2} матриці R_h . Розрахунок навантажень другого чинника матриці $A = \|a_{il}\|$. Повторюємо процедуру обчислень з матрицею перших залишкових коефіцієнтів кореляції R_l згідно з Крокami 3–5 даного алгоритму. Елементи другого власного вектора α_{i2} мають такі значення: $\alpha_{i2} = (-0,073998784; 0,477496356; 0,274396254; -0,358255367; 1; -0,165840708)$. Друге власне значення $\lambda_2 = \max(\beta_1^{(4)})$, $\lambda_2 = 0,421056$. Навантаження другого чинника відповідно рівні: $a_{i2} = (-0,0396763 \quad 45; 0,256021914; 0,147124587; -0,192087801; 0,53617564; -0,088919748)$.

Крок 7. Побудова матриці факторного відображення $A = \|a_{il}\|$, елементами якої є факторні навантаження, де $i = 1, \dots, n$ – параметри; $l = 1, \dots, r$ – фактори.

$$A = \|a_{ii}\| = \begin{pmatrix} 0,968013955 & -0,039676345 \\ 0,954376421 & 0,256021914 \\ 0,960972872 & 0,147124587 \\ -0,979212501 & -0,192087801 \\ -0,833657907 & 0,53617564 \\ 0,994430389 & -0,088919748 \end{pmatrix}$$

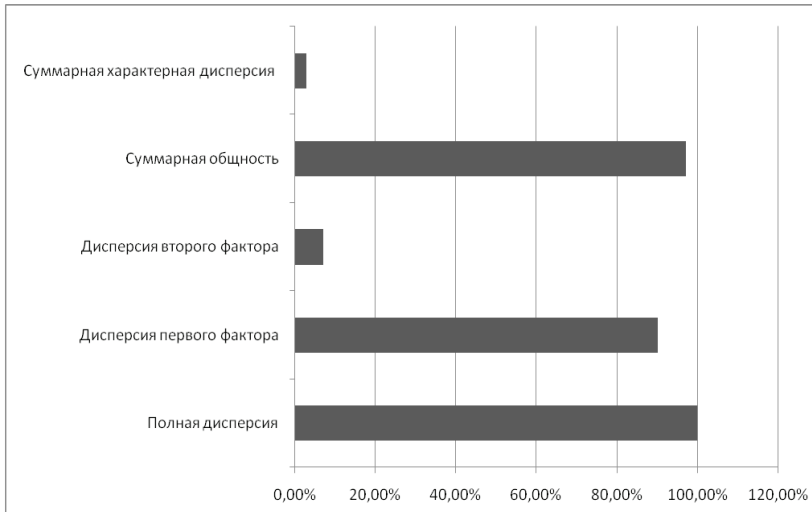
Матриця $A = \|a_{ii}\|$ дозволяє визначити, скільки загальних факторів необхідно для відображення всіх кореляцій між змінними, а також навантаження кожного фактора для всіх змінних і навантаження всіх факторів для однієї змінної.

Крок 8. Розрахунок квадратів факторних навантажень a_{ik}^2 і дисперсій кожного фактора. Обчислення часткою дисперсії факторів по матриці факторного відображення $A = \|a_{ii}\|$.

Повна дисперсія	6,00	100 %
Дисперсія першого фактора	5,414	90,23 %
Дисперсія другого фактора	0,421	7,02 %
Сумарна спільність	5,835	97,25 %
Сумарна характерна дисперсія	0,165	2,75 %

У результаті аналізу залишають у розгляді тільки ті фактори, які мають максимальні дисперсії і забезпечують 90–95 % від повної дисперсії факторів. Фактори, що мають дисперсію менше 5 % від повної дисперсії, з розгляду виключаються.

Крок 9. Визначення коефіцієнтів пріоритетності параметрів здійснюється на основі вкладу факторних навантажень кожної змінної в загальний чинник, що забезпечує максимум дисперсії від повної дисперсії факторів. На підставі проведених розрахунків (див. рис.) видно, що на два перших фактори доводиться 97,25 % від повної дисперсії і тільки 2,75 % припадає на сумарну характерну дисперсію, причому основна частина припадає на дисперсію першого фактора.



Розподіл часток дисперсії факторів, виражених у відсотках від повної дисперсії

Тому оцінку значимості параметрів y_i виконують залежно від факторних навантажень у перший загальний фактор (табл. 1).

Таблиця 1

Значення коефіцієнта пріоритетності параметрів

Параметр y_i	y_6	y_4	y_1	y_3	y_2	y_5
Коефіцієнт пріоритетності	0,988892	0,958857	0,937051	0,923469	0,910834	0,694986
Ранг	1	2	3	4	5	6

Крок 10. Розрахунок інформативності I отриманої сукупності факторів і втрат інформації при виключенні окремих факторів. Інформативність I сукупності двох перших факторів згідно з формулою

$$I = \frac{\lambda_1 + \lambda_2}{\sum_{i=1}^n \lambda_i}$$

склала 0,9725 од. інф. Втрати в інформативності при виключенні інших факторів складуть приблизно 2,75 % від повної інформа-

тивності. На підставі аналізу та розрахунків можна зробити висновок, що з усієї сукупності параметрів достатньо вибрати параметри з найбільшими коефіцієнтами пріоритетності, тобто y_3, y_4 для прогнозування технічного стану цифрових блоків комп'ютерної мережі в узагальненій контрольній точці. Інші параметри можна не враховувати, оскільки вони сильно корельовані між собою і мають менші вагові коефіцієнти. Запропонована методика формалізованого вибору параметрів дозволяє провести не тільки оцінку пріоритетності, побудувавши упорядковану послідовність параметрів у міру убування їх значущості за заданими критеріями, а й здійснити розрахунок інформативності факторів, виключити малозначні і малоінформативні фактори. З її допомогою можна здійснити науково обґрунтований вибір інформативною сукупністю параметрів для прогнозування інтенсивності руху через ДК, мінімізувати кількість параметрів для прогнозування без втрат в інформативності, зменшити розмірність масивів даних при обробці, а також скоротити витрати на їх контроль і прогноз. Для реалізації розрахунків і вибору параметрів за запропонованою методикою розроблена програма для ЕОМ в середовищі Delphi 6.

Проведено оцінку ефективності розробленого методичного забезпечення прогнозування інтенсивності руху, включаючи розроблену методику вибору параметрів і прогнозування. Для проведення порівняльної оцінки методик вибору параметрів або методик здійснення процесу прогнозування з розробленими методиками необхідно використовувати узагальнений показник, який включає в себе часткові показники параметрів або здійснення процесу прогнозування інтенсивності руху осіб або транспортних засобів. У загальному випадку узагальнений показник може бути представлений у вигляді функції

$$K = F(P_1, P_2, P_3, \dots, P_n), \quad (11)$$

де $P_1, P_2, P_3, \dots, P_n$ – часткові показники вибору параметрів.

Узагальнений показник має вигляд:

$$K = \sum_{i=1}^n P_i * W_i, \quad (12)$$

де W_1, W_2, \dots, W_n – вагові коефіцієнти.

При обчисленні узагальненого показника порівняльної оцінки необхідно враховувати значимість кожного часткового показника методик за допомогою вагових коефіцієнтів, які визначають методами експертних оцінок або за методикою аналізу ієрархій. При оцінці методик вибору параметрів прогнозу доцільно використовувати такі часткові показники.

Таблиця 2

Узагальнені показники методик вибору параметрів

№ з/п	Назва методик	Узагальнений показник		
		особи	транспортні засоби	вантажі
1	Методика вибору параметрів на основі інформаційного підходу	0,69	0,7	0,76
2	Методика вибору параметрів на основі методу аналізу ієрархій	0,62	0,53	0,71
3	Методика вибору параметрів на основі центроїдного методу факторного аналізу	0,68	0,72	0,75
4	Методика вибору параметрів на основі методу головних факторів	0,73	0,74	0,78

Як випливає з табл. 2, найбільш ефективною методикою для вибору параметрів прогнозу складових частин інтенсивності руху є методика на основі використання методу головних факторів.

Реалізація методичного забезпечення вибору параметрів приводить до підвищення достовірності прогнозування та зменшення економічних витрат.

Висновки. У статті була розроблена методика формалізованого вибору сукупності параметрів для прогнозування інтенсивності руху через ДК осіб, транспортних засобів і вантажів, що дозволяє провести не тільки оцінку пріоритетності з урахуванням їх взаємозв'язку, побудувавши упорядковану послідовність параметрів у міру убування їх значущості за заданими критеріями, а й здійснити розрахунок інфор-

мативності, вибрати найбільш значимі параметри, мінімізувати кількість параметрів без втрат в інформативності.

Дослідження показали, що даний підхід дає можливість обґрунтувати кількісний склад змін прикордонних нарядів у пунктах пропуску.

Перспективою подальших розвідок у даному напрямі є розробка методів пошуку вхідних параметрів моделі для прогнозу показників у діяльності ДПСУ з метою підвищення точності прогнозу.

Список використаної літератури

1. Кучков А. Ф., Лукашевич Н. Ф., Попов Г. П., Шумов В. В. Математическое моделирование служебно-боевых действий пограничных войск : учебник: в 3 томах. Москва : Академия ФПС России, 1997. Т. 1. 195 с. ; Т. 2. 191 с. ; Т. 3. – 240 с.
2. Бокс Дж., Дженкинс Г. Анализ временных рядов. Прогноз и управление. Москва : Мир, 1974.
3. Боровиков В. П., Ивченко Г. И. Прогнозирование в системе STATISTICA в среде Windows. Основы теории и интенсивная практика на компьютере. Москва : Финансы и статистика, 2000. С. 320.
4. Андрушук О. С. Прогнозування в охороні державного кордону на основі нейромережного підходу. *Збірник наукових праць Національної академії Державної прикордонної служби України імені Б. Хмельницького. № 54. Частина II. Хмельницький, 2010. С. 27–34.*

Рецензент – доктор технічних наук, професор Андрушук О. С.

Андрушко В. Рекомендации по применению методики прогнозирования при принятии решения начальником отдела пограничной службы

Представлены рекомендации начальнику отдела пограничной службы по принятию решений на основе прогнозирования с применением модели искусственной нейро-нечеткой сети по интенсивности деятельности пунктов пропуска Государственной пограничной службы Украины. Основные из них касаются работы с информацией (данными) и планирования смены пограничных нарядов с учетом результатов прогноза.

Исследования показали, что данный подход позволяет обосновывать количественный состав смен пограничных нарядов в пунктах пропуска.

Ключевые слова: прогноз, рекомендации, нейронная сеть, планирование.

Andrushko V. Recommendations for use method of forecasting when deciding head of the department border service

Management decision – one of the most important management processes. From its effectiveness depends largely on the success of the organization. Head of the Border Service of State Border and so must have technology justification, adoption, implementation, management decisions, without effective management of the organization in a complex operational (combat) situation almost impossible. One of the most important parts of developing management solutions are forecasting.

Developing prognosis is relatively complex and lengthy process. It consists of several stages:

1. Pre (research program).
2. Construction of the input (base) model intensity by means of system analysis.
3. Data collection forecast background.
4. Construction of time series of indicators.
5. Construction of a series of hypothetical (previous) search models.
6. Construction of hypothetical models of regulatory intensity.
7. Assessment of the validity, accuracy, validity (verification) forecast.
8. Developing recommendations for management decisions on the basis of a comparison search and normative models.
9. Analysis (examination) prepared forecasts and recommendations, as well as bring them considering the comments and report manager.
10. Before prognostic orientation based on a comparison of materials already produced new data on the forecast projected background and a new cycle of research.

Recommendations for planning based on forecasting. Closest prediction associated with planning. Plan and forecast – a mutually complementary

planning stage for the crucial role of the plan as a top-level management unit boundary.

The place of forecasting functions of border units. Based on the conventional concepts of control theory military units and law enforcement agencies, prediction plays a key role in the service of process management as a unifying link between objectively necessary functions of preparation and implementation of appropriate solutions.

Recommendations head of the department of border service decisions based on prediction models using artificial neural-fuzzy network on the intensity of checkpoints of the State Border Guard Service of Ukraine. The main concerns are related to the work of information (data) and planning border changes costumes based on the results of the forecast.

Keywords: *forecast, recommendations, neural network planning.*