

УДК 656.212

Р. В. ВЕРНИГОРА<sup>1\*</sup>, В. В. МАЛАШКИН<sup>2\*</sup>

<sup>1\*</sup> Каф. «Станции и узлы», Днепропетровский национальный университет железнодорожного транспорта имени академика В. Лазаряна, ул. Лазаряна, 2, 49010, Днепропетровск, Украина, тел. +38 (056) 373-15-12, эл. почта romav1@yandex.ua, ORCID 0000-0001-7618-4617

<sup>2\*</sup> Каф. «Станции и узлы», Днепропетровский национальный университет железнодорожного транспорта имени академика В. Лазаряна, ул. Лазаряна, 2, 49010, Днепропетровск, Украина, тел. +38 (056) 371-51-03, эл. почта malaxa79@mail.ru, .ORCID 0000-0002-5650-1571

## ОЦЕНКА УРОВНЯ ПРОФЕССИОНАЛЬНОЙ ПОДГОТОВКИ ДЕЖУРНЫХ ПО СТАНЦИИ С ИСПОЛЬЗОВАНИЕМ КОМПЬЮТЕРНЫХ ТРЕНАЖЕРОВ

Эффективность работы железнодорожных станций в значительной степени зависит от качества работы оперативно-диспетчерского персонала. В этой связи проблема качественной подготовки диспетчерского персонала для железнодорожного транспорта всегда была актуальной. Компьютерные тренажеры являются наиболее эффективным средством подготовки оперативно-диспетчерского персонала. В ДНУЗТ в течение последних лет ведется работа по созданию компьютерных тренажеров для подготовки дежурных по станции.

Одной из проблем, возникающих при разработке компьютерных тренажеров, является оценка качества профессиональной подготовки персонала по результатам тренировки. В статье предложено осуществлять оценку результатов работы на компьютерном тренажере методами дискриминантного анализа. При этом на основе вектора показателей о результатах тренировки работник относится к одной из групп, каждая из которых включает работников с определенным уровнем профессиональной подготовки.

На основе анализа были выбраны наиболее информативные показатели тренировки – простой поездов на станции, простой поездов на соседних станциях, простой поездов перед входным светофором и количество допущенных ошибок. Стандартизацию показателей предлагается осуществлять на максимум.

Начальную классификацию работников по результатам тренировки для получения исходной базы данных предлагается осуществлять методами кластерного анализа с учетом экспертных оценок. При этом наиболее целесообразно применять агрегативные алгоритмы постепенного объединения кластеров. Кроме того, в статье рассмотрена эффективность применения различных мер различия кластеров и стратегий их объединения.

Классификацию работника по результатам его тренировки предлагается осуществлять с использованием линейной дискриминантной функции Фишера, что обеспечивает наименьший уровень ошибки.

*Ключевые слова:* оперативно-диспетчерский персонал, профессиональная подготовка, компьютерный тренинг, оценка тренировки, дискриминантный анализ, кластерный анализ.

### Введение

Эффективность и безопасность функционирования железнодорожных станций зависит как от уровня их технического оснащения и технологии работы, так и от качества организации системы управления, основным звеном которой является оперативно-диспетчерский персонал (ОДП): дежурные по станции, маневровые диспетчера, дежурные по горке. Надежная и уверенная работа ОДП в любых условиях, особенно в нестандартных ситуациях, является важным фактором обеспечения безопасности движения и повышения эффективности работы железнодорожных станций. Вместе с тем, анализ причин браков, допускаемых на отечественных железных дорогах по хозяйству перевозок, показывает, что около 40 % случаев брака происходит по вине дежурных по станции и

поездных диспетчеров [1]. Подобная тенденция наблюдается и на зарубежных железных дорогах [2]. Неправильные или нерациональные управляющие действия дежурного по станции или маневрового диспетчера могут привести не только к снижению эффективности работы станции, но зачастую являются причиной возникновения аварийной ситуации [3]. В этой связи проблема профессионального отбора кадров и качественной подготовки работников оперативно-диспетчерского звена всегда являлась актуальной и требовала особого внимания.

Эффективным средством профессиональной подготовки диспетчерского персонала являются компьютерные тренажеры, которые, при существенно меньшей стоимости, по сравнению с аппаратными тренажерами, позволяют моделировать практически любую эксплуатационную ситуацию, а также решать широкий круг задач,

связанных с обучением ОДП. В последние годы компьютерный тренинг успешно применяется во многих странах мира практически на всех видах транспорта, в т. ч. и на железных дорогах [4].

#### **Анализ литературных источников и постановка задачи исследования**

Интенсивное развитие средств вычислительной техники и широкое внедрение информационных технологий на железнодорожном транспорте создало предпосылки к разработке программируемых тренажеров на базе серийных ПЭВМ. Особого внимания в этом отношении заслуживает опыт зарубежных железных дорог. Многие ведущие зарубежные железнодорожные компании широко используют компьютерный тренинг при подготовке ОДП [5-7]. С этой целью ими созданы специальные центры компьютерной подготовки персонала, выделяются значительные средства на разработку тренажеров. В США, Франции, Германии, Японии железнодорожные компании предписывают обязательный компьютерный тренинг для всех принимаемых на работу операторов и тренажерный курс переподготовки для всех работников оперативно-диспетчерского звена не реже одного раза в год. Более того, внутренние правила многих компаний предполагают обязательный восстановительный тренажерный курс после отпусков, болезни и т.д. [8, 9].

Учитывая актуальность проблемы профессионального обучения ОДП железных дорог, специалистами ДНУЖТ на протяжении нескольких последних лет ведется работа по созданию компьютерных тренажеров для подготовки дежурных по станции (ДСП) [10, 11]. В настоящее время разработаны и внедрены тренажеры для подготовки ДСП нескольких крупных сортировочных станций Украины, которые активно используются в вагонах-тренажерах, а также в нескольких профильных учебных заведениях Украины при подготовке студентов; кроме того, разработана технология подготовки ДСП станций железнодорожного участка с помощью тренажерного комплекса на базе локальной вычислительной сети [12].

Обычно основное внимание при создании тренажеров уделяется разработке имитатора объекта управления (станции), который включает его информационную модель и модель функционирования [13-15]. Однако, указанный имитатор без модели инструктора, обеспечивающей организацию, сопровождение и оценку тренировки, не может служить полноценным трена-

жером. При создании модели инструктора наибольшую трудность с методической точки зрения представляет выбор показателя для оценки проведенной тренировки, а также способа его расчета.

В существующих тренажерах оценка уровня профессиональной подготовки работника в большинстве случаев выполняется человеком-инструктором на основе анализа показателей выполненной тренировки, что снижает ее объективность [16, 17, 18]. В некоторых случаях предлагаются методики расчета обобщенного показателя для оценки работы на тренажере. При этом используется система начисления штрафных или премиальных баллов за те или иные действия тренируемого [19]. Однако, в связи с многообразием и сложностью решаемых в процессе работы на тренажере задач получение единого обобщенного показателя, который бы позволял объективно оценить деятельность ОДП, весьма затруднительно. Существуют тренажеры, в которых оценку выполненной тренировки предлагается производить на основе сравнения действий тренируемого с «эталонной» моделью деятельности ОДП, являющейся составной частью модели инструктора. При этом в процессе тренировки фиксируются все допущенные ошибки (несовпадения с «эталоном»), на основе чего определяется общая оценка тренажа [20]. Вместе с тем, формализация деятельности ОДП железнодорожного транспорта представляет собой весьма сложную задачу вследствие большого количества возможных вариантов действий работника в той или иной ситуации, а с учётом постоянно меняющейся оперативной обстановки разработка какого-либо «эталона» с целью использования его в тренажёре для оценки не представляется возможной.

В этой связи была поставлена задача разработать методику интегральной оценки уровня профессиональной готовности ОДП по результатам занятий на компьютерном тренажере.

#### **Общие принципы оценки результатов работы на компьютерном тренажере ДСП**

В процессе работы на тренажере все действия обучаемого фиксируются, а по окончании тренировки рассчитывается ряд показателей, характеризующих качество управления работой станции (средний простой поездов в парке, средний простой поездов на примыкающих линиях, количество допущенных ошибок, количество обработанных объектов и др.). На основе этих показателей формируется вектор класси-

фицирующих переменных тренируемого (объекта)  $X^* = \{x_1, x_2, \dots, x_N\}$ , где  $N$  – общее число показателей тренировки. Для решения поставленной задачи оценки предлагается использовать методы *дискриминантного анализа*, которые позволяют на основе значений вектора показателей  $X^*$ , отнести тренируемого к той или иной квалификационной группе. Каждая такая группа включает работников с определенным уровнем профессиональной подготовки.

Дискриминантный анализ [21, 22] является методом классификации с обучением, т.е. предполагает наличие определенной базы уже классифицированных по группам работников (обучающей выборки). Обучающая выборка представляет собой матрицу  $S$  размерностью  $M \times N$ , где  $M$  – объем выборки. Каждой строке матрицы (объекту выборки) поставлен в соответствие вектор  $X_i = \{x_{i1}, x_{i2}, \dots, x_{ij}, \dots, x_{iN}\}$  ( $i = 1 \dots M, j = 1 \dots N$ ), который включает значения показателей работы на тренажере для одного из классифицированных ранее работников и может быть представлен точкой в  $N$ -мерном пространстве. При этом каждый вектор  $X_i$  обучающей выборки отнесен к определенной квалификационной группе  $g_k$ , ( $g_k = 1 \dots G$ , где  $G$  – общее количество групп). Первоначальная классификация работников для получения обучающей выборки  $S$  может быть выполнена с помощью методов *кластерного анализа* [23] с учетом *экспертных оценок* [24].

#### **Выбор показателей для оценки уровня подготовки ОДП**

Выбор показателей, по которым выполняется оценка уровня подготовки работника по результатам его работы на тренажере, представляет собой весьма важную задачу. Как показывают исследования [21], с увеличением количества показателей  $N$ , используемых для классификации объектов, ее качество снижается. В этой связи среди множества показателей, которые определяются по результатам работы на тренажере ДСП, необходимо выделить подмножество наиболее значимых (информативных). Для этого следует определить показатели, значения которых значимо отличаются для групп с разным уровнем подготовки и которые играют определяющую роль при оценке уровня профессиональной обученности. Далее классификацию работника следует выполнять с учетом лишь этих показателей. При выборе информативных показателей следует также учи-

тывать мнение экспертов в области управления движением поездов и работой станций.

Среди показателей, получаемых в результате тренировки, с помощью дисперсионного анализа и метода множественных сравнений Шеффе [25] были выделены следующие информативные показатели: средняя продолжительность обслуживания поездов на станции –  $T_p$ , средний простой поездов на соседних станциях вследствие невозможности их приема –  $T_n$ , средний простой поездов перед входным светофором –  $T_c$ , суммарное число допущенных ошибок, приходящихся на один поезд –  $Z$ . Очевидно, что значимость ошибок, которые может допускать ДСП в процессе работы, может существенно отличаться. В этой связи был выполнен опрос работников ревизорского аппарата и оперативно-диспетчерского персонала Приднепровской железной дороги, по результатам которого была определена значимость (по 10-бальной шкале) каждой ошибки, которая может быть допущена ДСП в процессе работы.

#### **Стандартизация значений информативных показателей**

Показатели, используемые для классификации, измеряются в различных единицах и имеют разные диапазоны значений. В этом случае показатели с большим размахом значений будут оказывать доминирующее влияние на результаты классификации. Методы дискриминантного и кластерного анализа предполагают, что значения всех показателей классифицируемых объектов должны иметь одни и те же единицы измерения. Поэтому требуется стандартизация значений показателей как обучающей выборки  $S$ , так и вектора  $X^*$ , используемого при классификации конкретного объекта. Стандартизация предполагает переход к некоторому единообразному описанию данных и преобразование их в безразмерные величины. Выбор метода стандартизации представляет собой достаточно сложную задачу и существенно влияет на качество классификации. Это объясняется тем, что выбор метода может сильно изменять геометрию исходного  $N$ -мерного пространства. При разработке методики оценки результатов тренировки было выполнено исследование эффективности девяти различных методов стандартизации данных, предложенных в работах [22, 23]. Формулы для расчета конкретных значений по каждому из рассмотренных методов приведены в табл. 1.

Таблица 1

**Методы стандартизации значений при  
классификации объектов**

№	Формула	№	Формула	№	Формула
1	$X'_{i,j} = \frac{X_{i,j}}{R_j}$	4	$X'_{i,j} = \frac{X_{i,j}}{X_j^{\max}}$	7	$X'_{i,j} = \frac{X_{i,j} - X_j^{\min}}{\sigma_j^2}$
2	$X'_{i,j} = \frac{X_{i,j}}{\sigma_j^2}$	5	$X'_{i,j} = \frac{X_{i,j}}{\bar{X}_j}$	8	$X'_{i,j} = \frac{X_{i,j} - X_j^{\min}}{\sigma_j}$
3	$X'_{i,j} = \frac{X_{i,j}}{\sigma_j}$	6	$X'_{i,j} = \frac{X_{i,j} - X_j^{\min}}{\bar{X}_j}$	9	$X'_{i,j} = \frac{X_{i,j} - X_j^{\min}}{R_j}$

В табл. 1 приняты следующие обозначения:  $X_{i,j}$ ,  $X'_{i,j}$  – значение  $j$ -го показателя для  $i$ -го объекта обучающей выборки, соответственно, до и после стандартизации;  $X_j^{\min}$ ,  $X_j^{\max}$  – соответственно, минимальное и максимальное значение  $j$ -го показателя среди всех объектов выборки;  $\sigma_j^2$ ,  $\bar{X}_j$ ,  $R_j$  – соответственно, дисперсия, среднее значение и размах  $j$ -го показателя в обучающей выборке ( $R_j = X_j^{\max} - X_j^{\min}$ ).

Указанные методы стандартизации были исследованы в сочетании с различными методами кластерного и дискриминантного анализа при классификации результатов работы на тренажере контрольной группы из 50 человек, в которую вошли 15 профессиональных ДСП и 35 студентов 4-го курса факультета «Управление процессами перевозок» ДНУЖТ. При этом первоначально экспертным путем была выполнена «эталонная» классификация тренируемых на три группы: «высокий уровень подготовки» (группа А), «средний уровень» (группа Б) и «низкий уровень» (группа В). Эффективность каждого метода стандартизации оценивалась по двум критериям: вероятности ошибочной классификации –  $P_{\text{ош}}$  и величине суммарной внутригрупповой дисперсии –  $D_{\text{гр}}$ , которая определяет величину разброса значений отдельных показателей в квалификационных группах  $g_k$  после классификации. Критерий  $P_{\text{ош}}$  для каждого метода стандартизации определялся на основе сравнения результатов классификации, полученных с использованием данного метода, с «эталоном».

Для выполнения исследований были разработаны соответствующие программы, реализующие на ЭВМ различные алгоритмы дискриминантного и кластерного анализа. Исследования показали, что наилучшие результаты получены при стандартизации данных по методу 4 (табл. 1) – стандартизация на максимум. При этом вероятность ошибки  $P_{\text{ош}}$  в зависимости от метода классификации составила 3–5 %, а величина внутригрупповой дисперсии  $D_{\text{гр}}$  прини-

мала существенно меньшие значения, чем при использовании других методов стандартизации данных. Наихудшие результаты получены при стандартизации данных по методам 2 и 7, использование которых сильно искажает геометрию исходного  $N$ -мерного пространства, что обуславливает высокую степень вероятности ошибочной классификации 30–40 %.

**Получение обучающей выборки методами  
кластерного анализа**

Кластерный анализ представляет собой способ группировки (классификации) многомерных объектов, основанный на представлении результатов отдельных наблюдений точками подходящего геометрического пространства с последующим выделением групп (кластеров) этих точек (рис. 1). При этом каждый объект характеризуется вектором параметров (показателей)  $X = \{x_1, x_2, \dots, x_j, \dots, x_N\}$ . Задача методов кластерного анализа – выделить в исходных многомерных данных такие однородные пространства, чтобы объекты внутри группы были «похожи» в определённом смысле друг на друга, а объекты из разных групп – «не похожи». В данном случае под «похожестью» понимается близость объектов в многомерном пространстве. Кластерный анализ, в отличие от дискриминантного, представляет собой метод классификации без обучения [23].

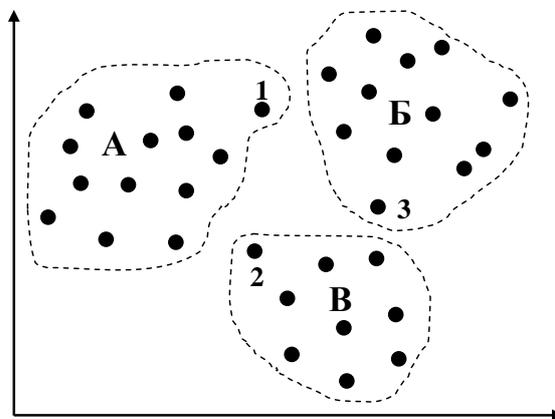


Рис. 1. Графическая интерпретация группировки объектов в кластеры

Методы кластерного анализа были использованы для получения исходной обучающей выборки  $S$ , которая является базой данных при оценке методами дискриминантного анализа уровня профессиональной подготовки конкретного работника по результатам его работы на тренажере ДСП. С этой целью был проведен ряд занятий на тренажере с контрольной группой из 50 человек с разным уровнем професси-

ональной подготовки.

Для решения задачи получения обучающей выборки  $S$  из множества существующих алгоритмов кластерного анализа [22, 23] были выбраны и исследованы алгоритмы прямой классификации, т.к. они ориентированы на выделение кластеров с заранее заданными свойствами, что соответствует условиям поставленной задачи. Кроме того, по сравнению с другими, алгоритмы прямой классификации имеют ряд преимуществ, к которым следует отнести простоту и ясность классификации, а также невысокую трудоёмкость. Алгоритмы прямой классификации можно разделить на *агломеративные* (стратегия объединения кластеров) и *дивизивные* (стратегия разделения).

В агломеративных алгоритмах каждый классифицируемый объект в начале классификации представляет собой отдельный кластер. На каждом шаге алгоритма происходит объединение двух наиболее близких кластеров. Это происходит до тех пор, пока число кластеров не достигнет определенного, в большинстве случаев, заранее заданного значения  $k$ . В дивизивных алгоритмах все объекты в начале классификации относятся к одному кластеру. Далее исходная совокупность постепенно разделяется до тех пор, пока не будет получена желаемая степень разделения (достигнуто заданное количество кластеров  $k$ ). Дивизивные алгоритмы целесообразно применять, когда исходную совокупность необходимо разделить на достаточно большое количество кластеров (более 10) [23], что не соответствует условиям задачи классификации работников по результатам работы на тренажере. Кроме того, дивизивные алгоритмы достаточно неустойчивы при классификации «спорных» объектов, находящихся на границах кластеров (объекты 1, 2, 3 на рис. 2). В этой связи, для выбора наиболее эффективного метода получения исходной обучающей выборки было выполнено исследование только агломеративных алгоритмов.

При классификации существенную роль играет также выбор меры различия между отдельными объектами. Мера различия между двумя объектами  $a$  и  $b$  представляет собой величину (расстояние)  $d_{ab}$ , которая тем больше, чем меньше сходство объектов  $a$  и  $b$ . При выборе метода классификации для получения обучающей выборки выполнено сравнение 4-х мер различия с целью определения наиболее подходящей для решения поставленной задачи [22, 23]. Формулы для расчета значений  $d_{ab}$  для каждой из рассмотренных мер различия приве-

дены в табл. 2. Здесь  $x_{aj}$ ,  $x_{bj}$  представляют собой значения  $j$ -го показателя для объектов  $a$  и  $b$ , соответственно ( $x_{aj} \in X_a$ ,  $x_{bj} \in X_b$ ,  $j = 1 \dots N$ ).

Таблица 2

**Меры различия при классификации объектов**

№	Мера	Формула
1	Эвклидова мера	$d_{ab} = \sqrt{\sum_j^N (x_{aj} - x_{bj})^2}$
2	Манхэттенская мера	$d_{ab} = \sum_j^N  x_{aj} - x_{bj} $
3	Мера Брея-Кертиса	$d_{ab} = \frac{\sum_j^N  x_{aj} - x_{bj} }{\sum_j^N x_{aj} + \sum_j^N x_{bj}}$
4	Канберровская мера	$d_{ab} = \sum_j^N \frac{ x_{aj} - x_{bj} }{ x_{aj}  +  x_{bj} }$

Для выбора наиболее эффективной меры различия с помощью ЭВМ был выполнен их сравнительный анализ при классификации данных о результатах тренировки контрольной группы. Эффективность каждой меры различия также оценивалась по показателям  $P_{\text{ош}}$  и  $D_{\text{тр}}$ . Исследование эффективности мер различия при формировании обучающей выборки методами кластерного анализа показало, что наилучшее качество классификации результатов работы на тренажере получено при использовании «канберровской» меры ( $P_{\text{ош}} = 4-6\%$ ). Данная мера различия также оказалась наиболее устойчива при классификации «спорных» объектов. Наихудшие результаты получены при расчете расстояний между объектами с использованием «манхэттенской» меры ( $P_{\text{ош}} = 25-30\%$ ).

Как отмечалось выше, агломеративные алгоритмы кластерного анализа предполагают постепенное объединение кластеров, пока их количество не достигнет заданного значения. На каждом шаге агломеративного алгоритма объединяются два кластера, расстояние между которыми минимально. При этом возникает проблема определения расстояний между отдельными кластерами. Способ расчета этого расстояния определяется принятой стратегией объединения.

Пусть имеются две группы объектов (кластеры)  $U$  и  $V$ , количество элементов в которых составляет  $n_u$  и  $n_v$ , соответственно; расстояние между этими кластерами  $d_{uv}$ . Допустим, что  $d_{uv}$  это минимальное среди всех возможных расстояний между всеми остальными кластерами.

В этом случае кластеры  $U$  и  $V$  объединяются в один кластер  $W$  с количеством элементов  $n_w = n_u + n_v$ . Рассмотрим некоторый кластер  $Y$ , который включает  $n_y$  элементов. Если перед объединением кластеров  $U$  и  $V$  известны расстояния  $d_{uy}$  и  $d_{vy}$ , то расстояние между новым кластером  $W$  и кластером  $Y$  составит:

$$d_{wy} = \alpha_u \cdot d_{uy} + \alpha_v \cdot d_{vy} + \beta \cdot d_{uv} + \gamma \cdot |d_{uy} - d_{vy}| \quad (1)$$

где  $\alpha_u$ ,  $\alpha_v$ ,  $\beta$ ,  $\gamma$  – параметры, определяющие сущность стратегии объединения.

При определении наиболее эффективного алгоритма классификации при получении обучающей выборки выполнено исследование 6-ти стратегий объединений [22]. В табл. 3 приведены параметры для определения расстояний  $d_{wy}$  (1) между отдельными кластерами для различных стратегий объединения.

Сравнительный анализ применения указанных стратегий для классификации результатов тренировок контрольной группы показал, что наиболее близкие результаты к «эталонной» обучающей выборке были получены при использовании «гибкой» стратегии объединения ( $P_{\text{ош}} = 4\text{--}6\%$ ). Худшие результаты получены при использовании стратегий «ближнего» и «дальнего» соседа ( $P_{\text{ош}} = 35\text{--}40\%$ ).

Таблица 3

**Параметры стратегий объединения в агломеративных алгоритмах классификации**

№ п/п	Наименование стратегии	Параметры стратегии
1	«Ближнего соседа»	$\alpha_u = \alpha_v = 0,5 \quad \beta = 0 \quad \gamma = -0,5$
2	«Дальнего соседа»	$\alpha_u = \alpha_v = 0,5 \quad \beta = 0 \quad \gamma = 0,5$
3	«Группового среднего»	$\alpha_u = n_u/n_w \quad \alpha_v = n_v/n_w$ $\beta = \gamma = 0$
4	«Центроидная»	$\alpha_u = n_u/n_w \quad \alpha_v = n_v/n_w$ $\beta = -\alpha_u \cdot \alpha_v \quad \gamma = 0$
5	«Квадратичная»	$\alpha_u = (n_u + n_y)/(n_w + n_y)$ $\alpha_v = (n_v + n_y)/(n_w + n_y)$ $\beta = -n_w/(n_w + n_y) \quad \gamma = 0$
6	«Гибкая»	$\alpha_u = \alpha_v = 0,625$ $\beta = -0,25 \quad \gamma = 0$

Таким образом, для получения обучающей выборки при классификации результатов работы на тренажере ДСП целесообразно использовать алгоритм кластерного анализа, основанный на «гибкой» стратегии объединения (табл. 3); при этом расстояния между отдельными объектами следует определять с помощью «канберровской» меры различия (табл. 2). Вместе с тем, обучающая выборка, полученная с помощью методов кластерного анализа, явля-

ется в определенной степени только рекомендацией, поскольку ее окончательное формирование необходимо осуществлять с учетом экспертных оценок соответствующих специалистов [24]. Следует отметить, что подобные выборки должны быть получены для различных начальных условий тренировки, которые в значительной степени определяют итоговые результаты тренажа. Применение разработанной методики позволяет автоматизировать и существенно ускорить процесс формирования подобных баз данных.

### Оценка результатов тренировки с помощью методов дискриминантного анализа

Для непосредственной оценки уровня профессиональной подготовки тренируемого на основе результатов его работы на тренажере в разработанной методике используется метод дискриминантного анализа. Сущность метода заключается в том, что вектор классифицирующих переменных  $\mathbf{X}^* = \{x_1, x_2, \dots, x_N\}$  тренируемого используется для классификации его к одной из групп ранее оцененных работников, данные о которых содержатся в базе данных тренажера. При этом оцениваемый работник относится к той квалификационной группе  $g_i$  ( $i=1 \dots G$ , где  $G$  – общее количество групп), для которой значение дискриминантной функции нормального распределения обращается в максимум [21, 22]:

$$f_i(\mathbf{X}^*) = (2\pi)^{-N/2} \cdot (\det \mathbf{S}_i)^{-1/2} \times$$

$$\times \exp\{-1/2 \cdot ((\mathbf{X}^* - \mathbf{m}_i) \cdot \mathbf{S}_i^{-1} \cdot (\mathbf{X}^* - \mathbf{m}_i)')\} \rightarrow \max \quad (2)$$

где  $\mathbf{m}_i$  – вектор математических ожиданий параметров  $x_1, x_2, \dots, x_N$  для  $i$ -й квалификационной группы;

$\mathbf{S}_i, \mathbf{S}_i^{-1}$  – соответственно, ковариационная и обратная ей матрицы размерностью  $N \times N$  для  $i$ -й группы.

Следует отметить, что аппроксимация данных в группах с помощью нормального закона является допущением, т.к. истинный закон распределения в большинстве случаев неизвестен. Вместе с тем, как показали исследования [22], аппроксимация неизвестного закона распределения нормальным дает вполне приемлемые результаты классификации. При этом истинные значения параметров неизвестного распределения математического ожидания  $\mu_i$  и ковариационной матрицы  $\Sigma_i$  заменяются их оценками  $\mathbf{m}_i$  и  $\mathbf{S}_i$ , вычисленными по данным  $i$ -й группы. Однако, с ростом объема обучающей выборки отклонение оценок  $\mathbf{m}_i$  и  $\mathbf{S}_i$  от  $\mu_i$  и  $\Sigma_i$  уменьшается, что

приводит к повышению качества классификации.

Как показали исследования, выполненные по результатам работы на тренажере ДСП группы студентов ДНУЖТ, методика оценки уровня их подготовки, основанная на (2), часто приводит к ошибочной классификации ( $P_{\text{ош}} = 20\text{--}25\%$ ). Кроме того, дискриминантная функция нормального распределения  $f_i(X^*)$  достаточно плохо классифицирует «спорные» объекты, находящиеся на границах групп.

В этой связи было выполнено исследование возможности использования для оценки результатов тренировки линейной дискриминантной функции Фишера [22, 23]. Как известно, если все группы  $g_i$ , среди которых выполняется классификация объекта  $X^*$ , имеют статистически одинаковую ковариационную матрицу  $S$ , то выражение (2) может быть представлено в виде линейной функции:

$$h_i(X^*) = (X^* \cdot \gamma_i) \cdot \lambda_i \rightarrow \max, \\ \text{где } \gamma_i = S^{-1} \cdot m_i, \lambda_i = 0,5 \cdot (m_i \cdot S^{-1} \cdot m_i') \quad (3)$$

Линейная функция Фишера имеет ряд преимуществ по сравнению с (2), так как вследствие своей линейности и большей надежности оценок ковариационных матриц даёт меньшую погрешность при классификации.

Один из методов дискриминантного анализа – *метод коалиций* [23] – предусматривает объединение тех групп  $g_i$ , у которых ковариационные матрицы статистически равны, в так называемые коалиции. Данные групп, входящих в одну коалицию, объединяются, и по ним вычисляется общая ковариационная матрица  $S$ . Далее классификация объекта  $X^*$  производится по максимуму линейной дискриминантной функции Фишера (3). Гипотеза о равенстве ковариационных матриц различных групп  $g_i$  проверяется по методу Бокса [23] или Барлетта-Андерсона [22, 25].

Выполненные исследования показали, что линейная функция Фишера при оценке результатов работы на тренажере ДСП показывает достаточно низкий уровень ошибочной классификации (5–7%) и более устойчива при классификации «спорных» объектов.

Как отмечалось выше, дискриминантный анализ является методом классификации с обучением. Это значит, что обучающая выборка может постоянно обновляться и увеличивать свой объем за счет данных об объектах, которые были правильно классифицированы. Такая организация также допускает возможность корректировки ошибочной классификации. Так, при выявлении подобной ошибки инструк-

тор, который руководит проведением занятий на тренажере, может откорректировать полученную тренируемым оценку и записать значения показателей его работы в нужную квалификационную группу в обучающей выборке. Таким образом, с ростом объема обучающей выборки вероятность ошибки при классификации уменьшается.

### Выводы

Необходимо отметить, что оценка профессиональных качеств работника по результатам его работы на тренажере, полученная с применением разработанной методики, не может и не должна рассматриваться как окончательная, а является лишь рекомендацией для лица, принимающего решение об уровне квалификации данного работника.

Приведенная методика оценки результатов тренировки была реализована в разработанных тренажерах ДСП и подтвердила свою эффективность при практическом использовании.

### БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Самсонкин, В. Г. Человеческий фактор в обеспечении безопасности железнодорожного транспорта / В. Г. Самсонкин // Инф-упр. системы на ж.д. тр. – 2000. – № 5–6. – С. 65–67.
2. Schmidt, A. Faktor mensch und sicherheit des grenzuberschreitenden schienenverkehrs / A. Schmidt, R. Miller // ETR: Eisenbahntechn. – 2004. – № 12. – P. 832-836.
3. Сидоренко Г. Г. Людський чинник як основа безпеки руху залізничного транспорту: аналітичний огляд / Г. Г. Сидоренко, О. А. Никифорова // Транспортні системи і технології перевезень : зб. наук. пр. Дніпропетр. нац. ун-ту заліз. трансп. ім. акад. В. Лазаряна – Дніпропетровськ, 2013. – Вип. 6. – С. 86-89.
4. Мямлин, С. В. Обзор конструкций технических средств для обучения специалистов железнодорожного транспорта / С. В. Мямлин, Е. П. Блохин, В. В. Жижко, Е. А. Письменный // Наука та прогрес транспорту. – 2006. – № 13. – С. 108-117.
5. Training now extends to all staff levels // International railway journal. – 1991. – № 4. – P. 22-23.
6. Тренажеры на железнодорожном транспорте // Железные дороги мира. – 1997. – № 4. – С. 31–34.
7. Kolmorgen, V.P., Naundorf, A. VIDINS – Videodatenbasierte interactive Schienenfahrzeug simulation / V. P. Kolmorgen, A. Naundorf // Eisenbahningenieur. – 2005. – № 6 – P. 40-44.
8. Человеческий фактор на железных дорогах // Железные дороги мира. – 1998. – № 5. – С. 32–38.
9. Braband, J. Kompetenzmanagement nach People CMM in der Eisenbahnindustrie / J. Braband // Signal+ Draht. – 2006. – № 3. – P. 6-8.
10. Бобровский, В. И. Повышение качества обучения оперативно-диспетчерского персонала желез-

нодорожних станцій с использованием компьютерных тренажеров / В. И. Бобровский, Р. В. Вернигора // Транспортні системи і технології : зб. наук. праць КУЕТТ. – Київ: КУЕТТ, 2003. – Вип. 3.– С. 54-61.

11. Вернигора, Р. В. Современные тренажерные системы для подготовки оперативно-диспетчерского персонала железнодорожных станций / Р. В. Вернигора // Вісник Східноукр. нац. ун-ту ім. В. Даля. – 2010. – № 1, ч. 2. – С. 34-40

12. Вернигора, Р. В. Підготовка ДСП станцій ділянки з використанням тренажерного комплексу / Р. В. Вернигора, В. В. Малашкін // Транспортні системи і технології перевезень : зб. наук. пр. Дніпропетр. нац. ун-ту залізн. трансп. ім. акад. В. Лазаряна – Дніпропетровськ, 2011. – Вип. 1. – С. 34-37

13. Бобровский, В. И. Эргатические модели железнодорожных станций / В. И. Бобровский, Козаченко Д. Н., Вернигора Р. В. // Транспортні системи і технології : зб. наук. праць КУЕТТ.– К.: КУЕТТ, 2004. – Вип. 5. – С. 80-86.

14. Бобровский, В. И. Количественная оценка технико-технологических параметров железнодорожных станций на основе эргатических моделей / В. И. Бобровский, Р. В. Вернигора, В. В. Малашкін // Наука та прогрес транспорту. – 2007. – № 16. – С. 50-57.

15. Bobrovskiy, V. I. Researching operation of the railway stations with using of their ergatic simulation models / V. I. Bobrovskiy, D. N. Kozachenko, R. V. Vernigora // Транспортні системи і технології перевезень : зб. наук. пр. Дніпропетр. нац. ун-ту залізн. трансп. ім. акад. В. Лазаряна – Дніпропетровськ, 2013.– Вип. 5 – С. 67-72.

16. Блохин, Е. П. Тренажер для обучения машинистов безопасным и экономичным способам вождения поездов / Е. П. Блохин //Залізничний транспорт України. – 1997. – № 2, – С. 25-28.

17. Сапунов, Н. А. Техническая учеба на технической станции / Н. А. Сапунов, А. И. Кожевников, С. И. Озеранский //Железнодорожный транспорт. – 2000. – № 9. – С. 75.

18. Тимергалин, Х. Н. Возможности автоматизированных обучающих систем и тренажеров / Х. Н. Тимергалин // Железнодорожный транспорт. – 2004. – № 4 – С. 84-87.

19. Кононов, В. А. Компьютерный тренажерный комплекс блочной маршрутно-релейной централизации / В. А. Кононов // Автоматика, связь, информатизация. – 2003 – № 4. – С. 28-30.

20. Титов, Е. В. Разработка тренажера для подготовки поездных диспетчеров : Автореф. дис. ... канд. техн. наук / Е. В. Титов – Москва : МИИТ, 1990. – 25 с.

21. Малиновский, Л. Г. Классификация объектов средствами дискриминантного анализа / Л. Г. Малиновский – Москва : Наука, 1979. – 180 с.

22. Энслейн К. Статические методы для ЭВМ / К. Энслейн. – Москва : Наука, 1986. – 464 с.

23. Мандель, И. А. Кластерный анализ / И. А. Мандель – Москва: Финансы и статистика, 1988. – 215 с.

24. Бешелев, С. Д. Математико-статистические методы экспертных оценок / С. Д. Бешелев, Ф. Г. Гурвич. – Москва : Статистика, 1974. – 160 с.

25. Андерсон, Т. В. Введение в многомерный статистический анализ / Т. В. Андерсон – Москва : Физматгиз, 1963. – 360 с.

*Стаття рекомендована до публікації д.т.н., проф. Бобровським В. І. (Україна)*

Надійшла до редколегії 03.10.2015.

Прийнята до друку 10.10.2015.

Р. В. ВЕРНИГОРА, В. В. МАЛАШКІН

## **ОЦЕНКА УРОВНЯ ПРОФЕССИОНАЛЬНОЙ ПОДГОТОВКИ ДЕЖУРНЫХ ПО СТАНЦИИ С ИСПОЛЬЗОВАНИЕМ КОМПЬЮТЕРНЫХ ТРЕНАЖЕРОВ**

Ефективність роботи залізничних станцій в значній мірі залежить від якості роботи оперативно-диспетчерського персоналу. В цьому зв'язку проблема якісної підготовки диспетчерського персоналу для залізничного транспорту завжди була актуальною. Комп'ютерні тренажери є найбільш ефективним засобом підготовки оперативно-диспетчерського персоналу. В ДНУЗТ протягом останніх років ведеться робота по створенню комп'ютерних тренажерів для підготовки чергових по станції.

Однією з проблем, що виникають при розробці комп'ютерних тренажерів, є оцінка якості професійної підготовки персоналу за результатами тренування. В статті запропоновано здійснювати оцінку результатів роботи на комп'ютерному тренажері методами дискримінантного аналізу. При цьому на основі вектора показників про результати тренування працівник відноситься до однієї з груп, кожна з яких включає працівників з певним рівнем професійної підготовки.

На основі аналізу було обрано найбільш інформативні показники тренування – простій поїздів на станції, простій поїздів на сусідніх станціях, простої поїздів перед входним світлофором та кількість допущених помилок. Стандартизацію показників пропонується здійснювати на максимум.

Початкову класифікацію працівників за результатами тренування для отримання вихідної бази даних пропонується здійснювати методами кластерного аналізу з врахуванням експертних оцінок. При цьому найбільш доцільно застосовувати агломеративні алгоритми поступового об'єднання кластерів. Окрім того, у статті розглянуто ефективність застосування різних мір розбіжності кластерів та стратегій їх об'єднання.

Класифікацію працівника за результатами його тренування пропонується здійснювати з використанням лінійної дискримінантної функції Фішера, що забезпечує найменший рівень помилки.

*Ключові слова:* оперативно-диспетчерський персонал, професійна підготовка, комп'ютерний тренінг, оцінка тренування, дискримінантний аналіз, кластерний аналіз.

R. VERNIGORA, V. MALASHKIN

## **ASSESSMENT TRAINING LEVEL OF RAILWAY DISPATCHERS USING COMPUTER SIMULATORS**

The efficiency of the railway stations largely depends on the quality of work of the dispatching personnel. In this context, the problem of quality of training supervisory staff for rail transport has always been important. Computer simulators are the most effective means of preparation of the dispatching personnel. In DNURT in recent years, the work on the creation of computer simulators for training the stations dispatcher.

One of the problems encountered in the development of computer simulators, is to assess the quality of training of staff on the results of the training. The paper proposed to assess the results of the computer simulator with using of discriminant analysis techniques. The employee on based indicators of the results of training belongs to one of the groups, each of which includes employees with a certain level of training.

On the basis of the analysis were selected the most informative indicators of training - a time of trains spent on station, on neighboring stations, before the input traffic lights and the number of errors. Standardizing encouraged to implement to the maximum of value.

Initial classification of the workers on the results of the training for the primary database are invited to carry out by the methods of cluster analysis and with using expert estimates. The most expedient to apply agglomerative algorithms gradual unification of clusters. In addition, the article examines the effectiveness of various measures differences clusters and their integration strategies.

The classification based on the results of its employee training is offered to implement use of linear discriminant function of Fisher, which provides the lowest level of error.

*Keywords:* operational and dispatching staff, professional training, computer training, evaluation of training, discriminant analysis, cluster analysis.