Санкт-Петербургский государственный университет **Факультет прикладной математики – процессов управления**

Марковская Вероника Александровна

Магистерская диссертация

Информационно-аналитический программный комплекс для государственного тестирования иностранных граждан по русскому языку

Направление 010300

Фундаментальные информатика и информационные технологии Магистерская программа «Автоматизация научных исследований»

Руководитель магистерской программы, доктор физ.-мат. наук, профессор Веремей Е. И. Научный руководитель,

кандидат физ.-мат. наук, доцент

Свиркин М. В.

Рецензент,

ст. преподаватель Севрюков С. Ю.

Санкт-Петербург 2015

Оглавление

Введение	. 3
Глава І. Миграционная политика Российской Федерации и государственное	
тестирование иностранных граждан	. 6
1. Государственное тестирование иностранных граждан по русскому	
языку	. 6
2. Информационно-логическая модель процесса тестирования	. 8
3. Анализ существующих подходов и комплексов тестирования	11
4. Цели и задачи	12
Глава II. Проектирование и математическое обеспечение программного	
комплекса проведения тестирования	13
1. Аппаратно-инструментальные средства	13
2. Программное обеспечение программного комплекса	13
3. Математическое обеспечение программного комплекса	15
4. Схема данных программного комплекса	20
Глава III. Реализация программного комплекса и проведение вероятностно-	-
статистического анализа результатов тестирования	24
1. Данные для проведения тестирования ЦЯТ	24
2. Реализация программного комплекса	27
3. Аналитические отчеты	42
4. Вероятностно-статистический анализ данных	48
Выводы	55
Заключение	57
	59
Приложение	31

Введение

Трудовая миграция – временная миграция экономически активного населения зарубежных стран в Россию с целью трудоустройства на ее территории.

В конце XX — начале XXI вв. среди миграционных потоков в Россию начинает заметно преобладать поток трудовой миграции, масштабы которого заметно выросли на протяжении последовавшего десятилетия. На данный момент по количеству мигрантов Россия занимает первое место в Европе и второе место в мире после США. По данным ФМС в России на данный момент находятся более 10 миллионов иностранных граждан. Наибольшее количество мигрантов находится в крупных городах России — Москве и Санкт-Петербурге, а также в областях — Московской и Ленинградской.

Согласно статистике Федеральной Миграционной Службы России [1] среднестатистический мигрант — это мужчина в возрасте от 18 до 29 лет, гражданин Узбекистана, приехавший для работы по найму на срок более года.

Общая статистика по целям пребывания на территории России иностранных граждан отличается для мужчин и женщин и представлена на рисунке 1.



Цель прибывания (женщины)

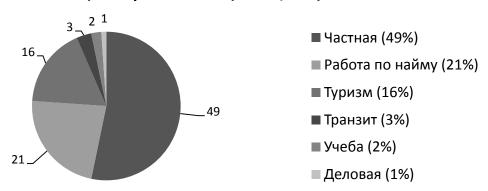


Рис. 1. Сведения в отношении иностранных граждан, находящихся на территории Российской Федерации (по целям прибывания)

По данным ФМС за 2008 г., структура трудовой миграции по сферам занятости сложилась следующим образом (в проц.) [2]:

Строительство	42,1
Обрабатывающее производство	9,9
Сельское хозяйство, охота	6,5
Транспорт и связь	3,9
Оптовая и розничная торговля, ремонт	16,9
Другие виды экономической деятельности	20,7

Табл.1. Структура трудовой миграции по сферам занятости по данным ФМС России за 2008 год

Основная проблема, связанная с иностранными гражданами, прибывающими на территорию Российской Федерации, — нелегальная миграция [3]. Решение этой проблемы — одно из актуальных направлений работы миграционных органов. Неурегулированный правовой статус иностранного гражданина наносит ущерб как самими мигрантам, так и

государству: рост криминальных проявлений, неисполнение налоговых обязательств; минимизация контактов с государственными структурами и незнание русского языка ведет к неспособности защитить свои права и интересы перед работодателем.

Соотношение легальной и нелегальной составляющей потока трудовой миграции в Россию существенно менялось на протяжении рассматриваемого десятилетия. Либерализация миграционного законодательства, изменение миграционного законодательства с 2006-2007 годов, основной целью которых являлось значительное сокращение нелегальной миграции и повышение уровня легализации временных трудовых мигрантов, сильно повлияли на это соотношение. Только за последнее время количество нелегальных мигрантов в стране уменьшилось на 1,3 миллиона человек. Заместитель главы ФМС России Е. Егорова отмечает [4], что во многом это произошло из-за введения экзамена по русскому языку и истории России, который послужил своеобразным тестом на адаптацию в современном российском обществе.

Можно с большой долей уверенности предположить, что в ближайшие годы, в связи с убылью трудоспособного населения, доля вакантных рабочих мест, на которые невозможно найти местного работника, будет расти и, соответственно, увеличится потребность работодателей в иностранной рабочей силе.

Глава I. Миграционная политика Российской Федерации и государственное тестирование иностранных граждан

1. Государственное тестирование иностранных граждан по русскому языку

Центр языкового тестирования СПбГУ успешно ведет свою деятельность с 1995 года, осуществляя государственное тестирование иностранных граждан по русскому языку, как иностранному. Иностранные граждане могут сдать экзамен для получения сертификата базового, элементарного, первого, второго, третьего и четвертого сертификационных уровней.

Приказом Министерства образования и науки РФ от 28 октября 2014 г. №1394 [5] Санкт-Петербургский государственный университет включен в перечень образовательных учреждений, проводящих экзамен по русскому языку как иностранному, истории России и основам законодательства Российской Федерации.

Комплексный экзамен проводится в целях подтверждения владения иностранными гражданами русским языком, знания истории России и основ законодательства Российской Федерации в соответствии с требованиями к минимальному уровню знаний, необходимых для сдачи указанного комплексного экзамена, утверждаемыми Министерством образования и науки Российской Федерации.

По своему содержанию данный тестовый экзамен является коммуникативным, так как в качестве объектов тестирования выступают общаться В ситуациях, умения иностранцев типичных требующих комплексного применения разных видов речевой деятельности (в различных сочетаниях), а также специфических коммуникативных умений, например, умения вступать в контакт (в устной или письменной форме), выбирать нужную тактику общения в зависимости от меняющихся условий общения и скорректировать её адекватно изменяющимся обстоятельствам.

Разделяют три категории комплексного экзамена:

- Комплексный экзамен для трудящихся мигрантов.
- Комплексный экзамен для лиц, желающих получить разрешение на временное проживание.
- Комплексный экзамен для лиц, желающих получить вид на жительство.

Данная работа посвящена созданию программного комплекса для центра языкового тестирования СПбГУ.

Часть «Русский язык» включает разделы по лексике, грамматике, письму, аудированию и говорению – в каждом из них у испытуемого должно быть не менее 60% правильных ответов. По остальным двум частям («Основы законодательства» и «История России») нужно верно ответить как минимум на половину вопросов. В целом экзамен оценивается по системе «зачет/незачет».

2. Информационно-логическая модель процесса тестирования

Иностранный гражданин или лицо без гражданства при подаче заявления на гражданство Российской Федерации должен представить документ, подтверждающий его владение русским языком на уровне, достаточном для общения в устной и письменной форме в условиях языковой среды.

С начала 2015 года по всей России комплексный экзамен уже сдали 440 тыс. мигрантов, 95% из которых получили соответствующие сертификаты [6].

Общая доля мигрантов в Петербурге, по мнению некоторых специалистов, неуклонно растет и к 2020 году может достигать сорока процентов [4].

Для ЦЯТ СПбГУ требуется разработать информационную систему для автоматизации анализа и учета данных о тестируемых, проведенных тестированиях и о выданных сертификатах.

Предметная область в данном случае – миграционная политика России и тестирование иностранных граждан по русскому языку. Выделяемые сущности – нормативно-правовая база, определяемая федеральными законами Российской Федерации; страны, граждане которых являются мигрантами России; уровни тестирования (цели прохождения тестирования, структура и иерархия экзамена); компоненты экзамена (предмет, составление заданий, определение минимальных проходных баллов и времени прохождения теста); тестируемый.

Основным объектом предметной области является «Тестируемый». Ограничимся определенным набором свойств объекта «Тестируемый». Перечислим эти свойства:

- 1) Фамилия Имя и Отчество
- 2) Дата рождения
- 3) Пол
- 4) Гражданство
- 5) Номер документа, удостоверяющего личность
- 6) Адрес, телефон (не обязательно)
- 7) Пройденные тестирования и результаты этих тестирований.
- 8) Выданные сертификаты и справки

Из общего описания предметной области известен ряд ограничений, существенных для процессов, которые относятся к рассматриваемой задаче. Предполагается, что тестируемый не может получить сертификат без успешной сдачи экзамена. Экзамен считается сданным успешно, если тестируемый набрал необходимые минимальные баллы по каждому компоненту экзамена. Минимальные баллы определяются министерством образования и науки РФ.

При неудовлетворительной сдаче одного или двух компонентов комплексного экзамена возможна пересдача только несданных компонентов. Если же число несданных компонентов более двух, то экзамен пересдается полностью.

Предполагается, что с данной информационной системой будут работать сотрудники центра тестирования, и они должны иметь возможность решать с помощью нее следующие задачи:

- 1) Ввод и редактирование данных о тестируемом.
- 2) Ввод результатов проведенного тестирования.
- 3) Печать сертификатов, подтверждающих успешное прохождение государственного тестирования по русскому языку.
- 4) Печать справок с результатами тестирования.
- 5) Печать протоколов печати и реестра выдачи сертификатов.

При успешном прохождении экзамена тестируемый получает сертификат. Срок действия сертификата зависит от уровня тестирования. Для комплексных экзаменов срок действия сертификата составляет 5 лет, для уровней ТРКИ – сертификат выдается бессрочно.

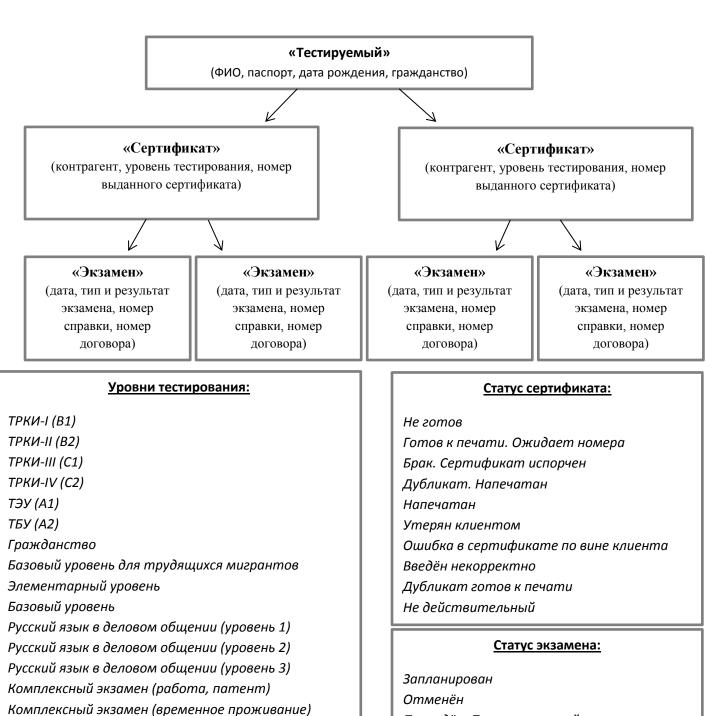


Рис. 2. Сущности предметной области

Комплексный экзамен (вид на жительство)

Проведён. Положительный результат

Проведён. Отрицательный результат

3. Анализ существующих подходов и комплексов тестирования

Подавляющее большинство центров языкового тестирования, находящихся на территории РФ, проводят тестирование в письменном виде.

Однако в настоящее время существуют программные комплексы по проведению электронного тестирования по русскому языку для иностранных граждан. Преимуществами таких программных комплексов являются: оперативность при подведении итогов тестирования, хранение и учет данных в единой базе, возможность подсчета необходимых статистик по заданным параметрам.

Наиболее известный программный комплекс по проведению электронного тестирования иностранных граждан — программа «Эльтерус» [7]. Она была разработана в июле 2012 года совместно со специалистами из СПбГУ и при научно-методической поддержке Международной ассоциации преподавателей русского языка и литературы. Впервые программный комплекс «Эльтерус» стал использоваться в Едином Миграционном Центре в Санкт-Петербурге.

Однако в таких программных комплексах нет модуля аналитической обработки информации или модуля поддержки принятия оптимальных решений на основе результатов моделирования и прогнозирования, а также систем создания аналитических отчетов.

Внедрение механизмов электронного тестирования и оценки знаний позволит внедрить дифференцированный подход и сформировать качественную статистику владения Русским языком, для того чтобы в последствии реализовать эффективные инструменты в области миграционной политики.

4. Цели и задачи

Цель настоящей работы заключается в разработке информационноаналитического программного комплекса для Центра Языкового Тестирования СПбГУ с возможностью проведения вероятностностатистического анализа и поддержки принятия решений в области миграционной политики.

Основные задачи исследования:

- анализ предметной области с выделением сущностей;
- построение информационно-логической модели
- проектирование и реализация программного комплекса;
- проведение вероятностно-статистического анализа и прогнозирования;
- создание системы аналитической отчетности;
- создание модуля поддержки принятия решений.

Разрабатываемый программно-аналитический комплекс должен обеспечивать следующий функционал:

- ведение справочников;
- автоматизация ввода и вывода информации о тестируемых;
- проверка вводимых данных на корректность;
- сопровождение процесса тестирования;
- печать и учет сертификатов, подтверждающих знание русского языка;
- создание автоматических отчетов для планового финансового отдела и бухгалтерии.

Глава II. Проектирование и математическое обеспечение программного комплекса проведения тестирования

1. Аппаратно-инструментальные средства

Программный комплекс «База сертификатов по русскому языку» предполагается для использования сотрудниками Центра Языкового Тестирования. Приложение находится на удаленном рабочем столе, для доступа к которому необходимо иметь единую учетную запись СПбГУ. Использование единой учетной записи обладает рядом преимуществ, таких как повышенная безопасность, разграничение прав доступа к тем или иным функциям, контроль ввода данных в систему и другие.

Основные данные, необходимые для работы приложения хранятся на сервере, который расположен на территории главного здания СПбГУ.

Таким образом, для доступа к приложению необходимо иметь персональный компьютер с выходом в интернет и внутреннюю сеть СПбГУ, единую учетную запись СПбГУ с необходимыми правами доступа. На персональный компьютер должна быть установлена операционная система, поддерживающая сеансы удаленного рабочего стола.

Для доступа к сайту Центра Языкового Тестирования необходимо иметь персональный компьютер с выходом в интернет и браузер с поддержкой скриптов JavaScript и включенными cookies.

2. Программное обеспечение программного комплекса

Для разработки программного комплекса был выбран язык программирования Microsoft Visual C#, экспресс версия которого

Microsoft бесплатно. предоставляется компанией Данный язык программирования отличается от остальных тем, что на нем можно реализовать необходимый графический интерфейс гораздо быстрее, чем на других языках. Также основным достоинством языка является способность подключения библиотек, c созданных помощью других языков программирования.

При создании приложения были использованы как стандартные, так и разработанные сотрудниками УСИТ СПбГУ библиотеки. К стандартным библиотекам, задействованным в проекте, относятся:

- Microsoft.Office.Interop.Word и Microsoft.Office.Interop.Excel (реализуют взаимодействие с объектными моделями Word и Excel) [8],
- System.Data.Entity (обеспечивает доступ к ключевым функциям платформы Entity Framework) [9],
- System.Xml.Linq (предлагает удобные шаблоны для выполнения запросов и обновления данных) [10].

В проекте также используются две дополнительные библиотеки:

- iTextSharp (свободно распространяемая библиотека, позволяющая генерировать PDF-документы) [11],
- WordOut (включающая в себя дополнительные функции по работе с текстовыми файлами форматов doc, excel, pdf).

Для хранения данных используется СУБД Microsoft SQL Server.

В качестве технологии разработки сайта для Центра Языкового тестирования была выбрана ASP .NET [12]. Основными преимуществами платформы ASP .NET MVC являются:

• Разделение приложения на модель, представление и контроллер облегчает управление сложными структурами приложения.

- Полноценный язык. Платформа позволяет использовать любой .NET совместимый язык (наиболее популярным является С#).
- Компиляция программ. Компилируемый код выполняется быстрее, большинство ошибок отлавливается еще на стадии разработки.

Статистическое изучение данных тестирования ведется в двух основных направлениях: статике и динамике. Под статикой будем понимать изучение данных по ряду признаков: пол, возраст и национальность тестируемых, минимальный, средний и максимальный возраст тестируемых, средние баллы по компонентам экзамена, количество попыток сдачи экзамена, цель прохождения тестирования и другие. Под динамикой будем понимать изменения количества тестируемых в месяц, изменение средних показателей с течением времени и пр.

3. Математическое обеспечение программного комплекса

Ввиду нового федерального закона «Об утверждении формы, порядка проведения экзамена по русскому языку как иностранному, истории России и основам законодательства Российской Федерации, требований к минимальному уровню знаний, необходимых для сдачи указанного экзамена» [13] количество тестируемых в Центре Языкового Тестирования резко возросло. Если среднее число тестируемых в месяц в 2014 году равнялось 1200, то количество тестируемых в январе превысило 13000 человек, а в феврале — 25000.

Резкие перепады значений временного ряда ведут к невозможности применения стандартных методов прогнозирования к исходным данным. В качестве компонент временного ряда будем рассматривать отношение количества тестируемых за месяц к среднему годовому количеству тестируемых.

Очевидно, количество тестируемых есть дискретная случайная величина, присутствует сезонная компонента. Для анализа временных рядов чаще всего используют две модели: аддитивную и мультипликативную [14].

В аддитивной модели временной ряд представляется в виде суммы компонент и в общем виде выглядит следующим образом:

$$y_t = b_t + f_t + c_t + e_t,$$

где y_t — фактическое значение уровня ряда на момент времени t, b_t — значение тренда, f_t — значение фактора сезонности, c_t — циклическая компонента, e_t — ошибка.

Мультипликативная модель предполагает, что каждый уровень временного ряда представляется в виде произведения компонент (при тех же обозначениях):

$$y_t = b_t \cdot f_t \cdot c_t \cdot e_t$$
.

Для прогнозирования временного ряда с трендом и сезонностью будем использовать трехпараметрическую модель Хольта-Винтерса [14, 15], которая построена на основе объединения двухпараметрической модели Хольта и сезонной модели Винтерса:

$$\begin{cases} S_{t} = \alpha \cdot \frac{y_{t}}{f_{t-T}} + (1 - \alpha) \cdot (S_{t-1} + b_{t-1}); \\ f_{t} = \beta \cdot \frac{y_{t}}{S_{t}} + (1 - \alpha) \cdot f_{t-T}; \\ b_{t} = \gamma \cdot (S_{t} - S_{t-T}) + (1 - \gamma) \cdot b_{t-1}, \end{cases}$$

где

t — независимая переменная (время),

T — количество шагов временного ряда, определяющих периодичность сезонности (T=12),

 y_t – фактический уровень ряда на момент времени t,

 S_t — сглаженное значение фактического уровня y_t ,

 f_t – сглаженное значение фактора сезонности,

 b_t - сглаженное значение тренда,

 α , β , γ — параметры сглаживания (α , β , $\gamma \in [0;1]$).

Прогноз на к-периодов строится по формуле:

$$y_{t+k} = (S_t + k \cdot b_t) \cdot f_{t+k-T}$$

Параметры сглаживания подбираются в интервале от 0 до 1 с шагом 0,1.

После построения любой модели процесса возникает необходимость оценить соответствие модели реальной ситуации. Для прогнозирования каких-либо показателей пригодны лишь качественные модели, адекватно отображающие истинные зависимости и тенденции.

В основе верификации модели (т.е. проверки ее адекватности) лежит исследование остатков, которые определяются в виде разностей:

$$e_t = y_t - \hat{y}_t,$$

где \hat{y}_t — систематическая составляющая временного ряда с уровнями y_t . Полученная последовательность остатков $e_1, e_2, \dots e_n$ есть последовательность реализаций некой случайной величины ε_t , системы случайных величин $\varepsilon_1, \varepsilon_2, \dots \varepsilon_n$.

Для верификации модели необходимо выполнение следующих условий:

- случайные величины $\varepsilon_1, \varepsilon_2, \dots \varepsilon_n$ независимы от времени t;
- все случайные величины ε_t являются гауссовскими с нулевым средним и постоянной дисперсией;

• отсутствие автокорреляции между случайными величинами ε_t и $\varepsilon_{t+k}, k \geq 1.$

Первое условие можно проверить различными критериями, для конкретизации рассмотрим критерий Аббе, также называемый критерием квадратов последовательности разностей.

Проверяется нулевая гипотеза:

 H_0 : элементы выборки ε_1 , ε_2 , . . . ε_n независимы.

Конкурирующая гипотеза:

 H_1 : элементы выборки ε_1 , ε_2 , . . . ε_n зависимы.

Статистическим критерием проверки гипотезы H_0 служит статистика $\gamma(n) = \frac{q^2(n)}{\varsigma^2(n)}$, где

$$q^{2}(n) = \frac{1}{2(n-1)} \sum_{i=1}^{n-1} (x_{i+1} - x_{i})^{2},$$

$$S^{2}(n) = \frac{1}{n-1} \sum_{i=1}^{n} (x_{i} - \bar{x})^{2}.$$

Для формирования области принятия гипотезы H_0 и критической области определяется критическая точка $\gamma_{\rm kp}(n)$, которая кроме п также зависит от выбранного уровня значимости α . Значение критической точки при $n \le 60$ можно найти по соответствующим таблицам.

Нулевая гипотеза о независимости элементов принимается в случае: $\gamma(n) \, > \gamma_{\rm kp}(n) \text{, и отвергается в противном случае.}$

Статистический критерий Дарбина-Уотсона d проверки нулевой гипотезы H_0 : $r_1=0$ выражается через остатки в виде соотношения сумм:

$$d = \frac{\sum_{t=2}^{n} (\varepsilon_t - \varepsilon_{t-1})^2}{\sum_{t=1}^{n} \varepsilon_t^2}.$$

Данное отношение легко преобразуется к виду:

$$d = 2\left(1 - \frac{\sum_{t=2}^{n} \varepsilon_{t} \varepsilon_{t-1}}{\sum_{t=1}^{n} \varepsilon_{t}^{2}}\right) - \frac{e_{1}^{2} + e_{n}^{2}}{\sum_{t=1}^{n} \varepsilon_{t}^{2}} \approx 2(1 - r_{1})$$

Если корреляция отсутствует, то значение d близко к 2. Сравнивая вычисленное значение d с соответствующими табличными интервалами, можно сделать вывод о наличии или отсутствии автокорреляции.

Наряду с адекватностью качество построенной модели определяется ее точностью. Точность модели характеризуется множеством случайных отклонений от прогнозируемых значений.

Характеристики различия между наблюдаемыми и прогнозируемыми значениями называют ошибкой прогноза. Ошибки прогноза классифицируются как абсолютные и относительные. Чем меньше эти ошибки, тем качественнее модель.

Абсолютной ошибкой прогноза называют разность между прогнозным и наблюдаемым значениями показателя:

$$\Delta_t = \hat{y}_t - y_t.$$

Относительной ошибкой прогноза называют отношение абсолютной ошибки прогноза к наблюдаемому значению:

$$\delta_t = \frac{\Delta_t}{y_t} = \frac{\hat{y}_t - y_t}{y_t}.$$

Приведем пример средних ошибок, характеризующих всю совокупность в целом.

Средняя квадратическая ошибка прогноза:

$$\bar{S} = \sqrt{\frac{1}{n} \sum_{t=1}^{n} (\hat{y}_t - y_t)^2}.$$

Относительная средняя квадратическая ошибка прогноза:

$$\bar{\gamma} = \sqrt{\frac{1}{n} \sum_{t=1}^{n} \left(\frac{\hat{y}_t - y_t}{y_t}\right)^2}.$$

4. Схема данных программного комплекса

Данные размещаются в БД «Тестирование», которую можно логически разделить на две части. Первая и основная часть - данные, необходимые для работы приложения:

- персональные данные тестируемого (ФИО, дата рождения, гражданство, серия и номер паспорта и пр.);
- результаты тестирования (уровень тестирования, баллы по компонентам экзамена, дата экзамена, результат экзамена);
- данные о выданных сертификатах (номер бланка, регистрационный номер, дата выдачи);
- справочники (список стран, список уровней тестирования, список типов договоров, список контрагентов, список компонентов экзамена и пр.);
- прочие данные (данные о заключенных договорах, стоимости тестирования, переменные для печати).

База данных состоит из множества связанных таблиц, каждая из которых содержит сведения об объектах определенного вида. Основные таблицы (рис. 3), представляющие интерес для анализа данных, это: «Тестируемый», «Сертификат», «Экзамен», «Оценки», «Страны».

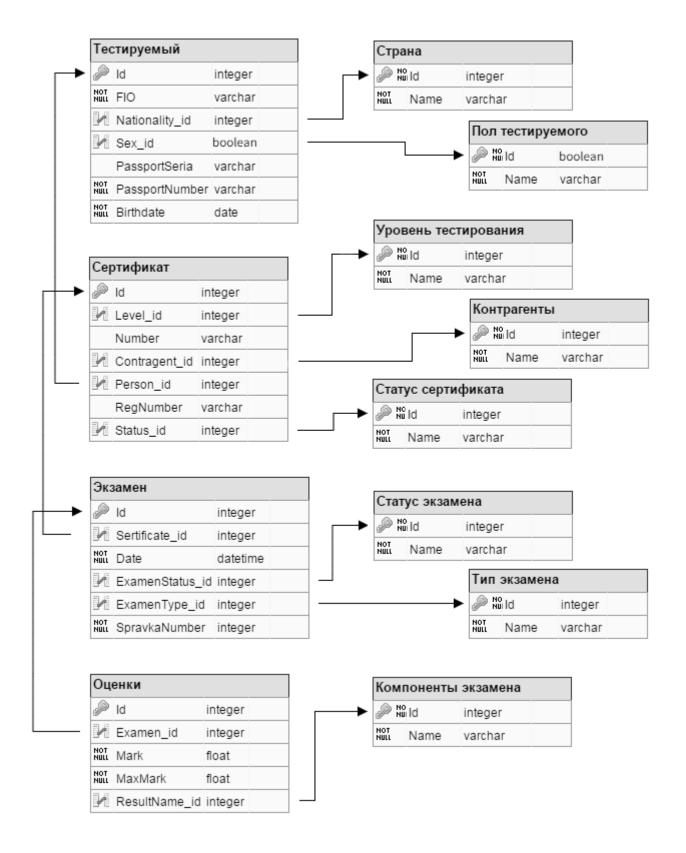


Рис. 3. Логическая схема базы данных «Тестирование» (данные, необходимые для работы с приложением)

Вторая часть – данные, необходимые для работы с сайтом (рис. 4):

- Учетные записи пользователей (логин, пароль, ФИО сотрудника и пр.)
- Данные организаций, которым предоставлен доступ к системе.

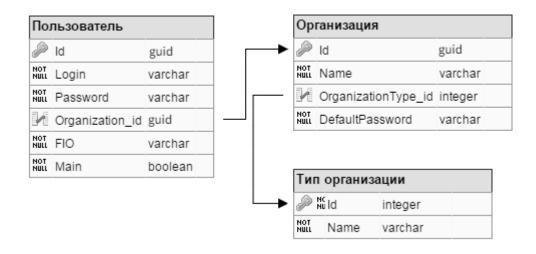


Рис 4. Логическая схема данных «Тестирование» (данные, необходимые для работы с сайтом)

Предполагается разделение на три основных типа учетных записей в зависимости от необходимой функциональности:

- «Администратор сайта»: создание, редактирование и удаление аккаунтов для организаций.
- «Сотрудник УФМС»: проверка номеров сертификатов на их наличие в базе ЦЯТ и корректности указанных данных.
- «Сотрудник центра тестирования»: загрузка результатов тестирования иностранных граждан в виде xls-файла (csv-файла) или с помощью заполнения специальной формы (рис. 5).

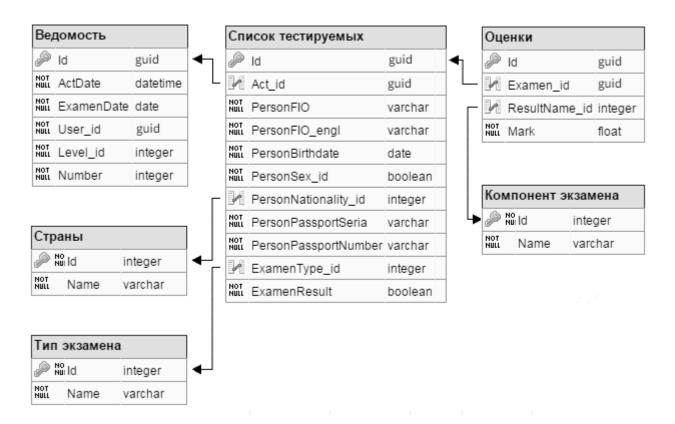


Рис. 5. Схема размещения ведомости и результатов тестирования в базе данных

Глава III. Реализация программного комплекса и проведение вероятностно-статистического анализа результатов тестирования

1. Данные для проведения тестирования ЦЯТ

За 2013-2015 года в Центре Языкового Тестирования более 120 тысяч мигрантов получили возможность сдать экзамен на знание русского языка. Это представители более ста стран, мужчины и женщины разных возрастов.

Большая часть тестируемых приходится на жителей Республики Узбекистан (более 60 тыс. человек) и жителей Республики Таджикистан (более 22 тыс. человек), далее следуют жители Украины (12 тыс. человек), Республики Молдовы (7 тыс. человек), Киргизии (3 тыс. человек), Республики Азербайджан (3 тыс. человек), Республики Армения (2 тыс. человек). Завершают список Китай, Республика Корея и Казахстан – количество тестируемых из этих стран составляет 1-1,5 тыс. человек.

Иностранным гражданам предоставляется возможность сдать экзамен на знание русского языка, и в зависимости от цели сдачи экзамена можно выбрать:

- ТРКИ-I (B1)
- ТРКИ-II (B2)
- ТРКИ-III (C1)
- ТРКИ-IV (C2)
- ТЭУ (A1)
- ТБУ (A2)
- Гражданство
- Базовый уровень для трудящихся мигрантов
- Элементарный уровень
- Базовый уровень

- Комплексный экзамен ИР (работа, патент)
- Комплексный экзамен РВ (временное проживание)
- Комплексный экзамен ВЖ (вид на жительство)

До федерального закона («О правовом положении иностранных граждан в Российской Федерации») большая часть тестируемых сдавала экзамен для оформления разрешения на работу в Российской Федерации (уровень - «Базовый уровень для трудящихся мигрантов»). На втором месте — экзамен, подтверждающий знание русского языка на уровне, достаточном для получения гражданства Российской Федерации (уровень — «Гражданство»).

Однако с 1 января 2015 года экзамен «Базовый уровень для трудящихся мигрантов» не проводится, вместо него появилось три новых уровня тестирования: комплексный экзамен – ИР (работа, патент), комплексный экзамен – РВ (временное проживание), комплексный экзамен – ВЖ (вид на жительство). Все три уровня комплексного экзамена являются наиболее востребованными. С января этого года комплексный экзамен для получения разрешения на работу сдало более 90 тыс. человек, для разрешения на временное проживание – более 10 тыс. человек, для вида на жительство – более 5 тыс. человек.

Необходимо отметить, что более 90% иностранных граждан успешно сдают тестирование. Наиболее высокие оценки по компонентам экзамена получают граждане бывших Советских Республик. Это обусловлено тем, что на территории этих стран наблюдается широкое использование русского языка в повседневной жизни. Например, средний балл для тестируемых из Украины составляет более 90-95% от максимального количества баллов, в то время как для тестируемых из Польши этот показатель равен 75%, для Словакии — 80%. Однако, несмотря на то, что некоторые страны являются бывшими Советскими Республиками, уровень сдачи компонентов экзамена

хоть и удовлетворительный, но не является высоким. Таким образом, для тестируемых из Узбекистана и Таджикистана средний балл составляет 75-80% от максимального (Табл. 2).

	Кол-во	Компоненты экзамена						
Страна	тестируемых	Пс	Гв	Гр	q_{T}	Ауд	Ист	Зак
Все страны	119669	80 %	78%	81 %	84 %	79%	76 %	76 %
Республика Узбекистан	63603	78 %	74 %	78 %	83 %	78 %	74 %	74 %
Республика Таджикистан	22767	77 %	75 %	78 %	80 %	76 %	74 %	75 %
Украина	13167	91 %	94 %	98 %	94 %	90 %	86 %	86 %
Республика Молдова	7248	85 %	90 %	94 %	92 %	87 %	83 %	86 %
Республика Азербайджан	3285	80 %	81 %	82 %	84 %	81 %	74 %	74 %
Кыргызстан	3284	82 %	82 %	82 %	85 %	82 %	74 %	75 %
Республика Армения	1743	84 %	90 %	86 %	89 %	85 %	78 %	79 %

Табл. 2. Средние баллы по компонентам экзамена для стран с наибольшим количеством тестируемых (Пс – письмо, Гв – говорение, Гр – Грамматика, Чт – чтение, Ауд – аудирование, Ист – история РФ, Зак – основы законодательства РФ)

2. Реализация программного комплекса

Информационная система реализуется в среде программирования С# и .Net и представляет собой приложение с графическим интерфейсом, состоящее из экранных форм, диалоговых окон и других элементов управления.

Главное меню программы состоит из следующих пунктов:

• Тестируемые:

- о Добавление нового тестируемого
- о Список тестируемых
- о Ввод оценок за экзамен
- о Объединение карточек тестируемых

• Загрузка данных:

- о загрузка данных в виде csv-файла;
- о загрузка данных в виде xls-файла;
- о создание выгрузки в базу РУДН;
- о заявка на дубликаты.

• Печать сертификатов:

- о печать сертификатов;
- о печать справок;
- о изменение статусов сертификатов.

• Протоколы:

о список протоколов.

• Справочники:

- о максимальные баллы;
- о список стран;
- о список контрагентов;
- о переменные для печати.

• Отчетность:

- о формирование ответа на запрос УФМС;
- о реестр заключенных договоров;
- о отчет по проведенному тестированию;
- о отчетность по контрагентам.
- Открыть временную папку.
- Статистика:
 - о полная статистика;
 - о результаты тестирования;
 - о отчет по форме-1;
 - о отчет по форме-2;
 - о отчет по форме-3.

Главная форма системы содержит строку выпадающих меню, команды которых обеспечивают работу с системой. В процессе работы из главной формы вызываются подчиненные формы для ввода и редактирования исходных данных, просмотра данных и результатов обработки.

Карточка тестируемого

В карточке тестируемого (Рис. 6) отображаются его персональные данные: ФИО, ФИО латиницей, дата рождения, гражданство, родной язык, тип паспорта (паспорт или вид на жительство), номер и серия паспорта. В правой части карточки есть возможность записать тестируемого на экзамен, указав уровень тестирования, выбрав дату экзамена и заключив договор.

Во второй закладке «Сведения об экзаменах» (Рис. 7) содержится информация о тестированиях.

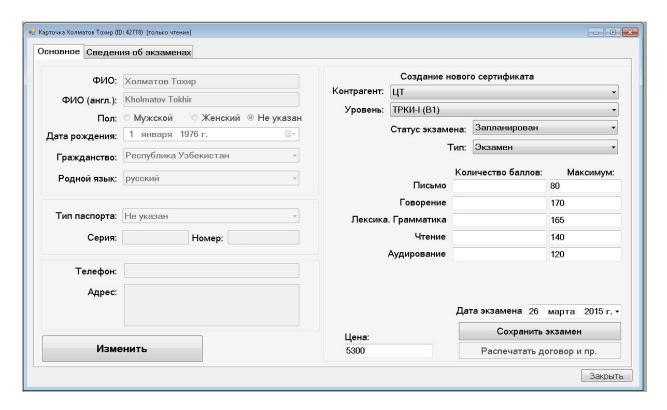


Рис. 6. Карточка тестируемого. Основные сведения

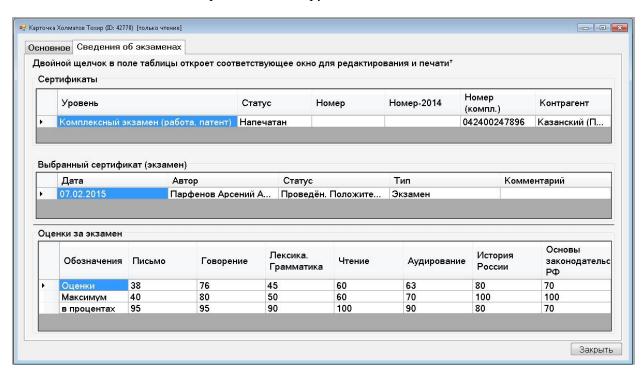


Рис. 7. Карточка тестируемого. Сведения о результатах экзаменов

Список тестируемых

Предусмотрена выборка тестируемых по различным критериям: гражданство тестируемого, наименование контрагента, результат экзамена, дата экзамена, уровень тестирования, статус сертификата, номер сертификата; и три варианта сортировки тестируемых — по ФИО тестируемого, по номеру сертификата и номеру сертификата комплексного экзамена (Рис. 8).

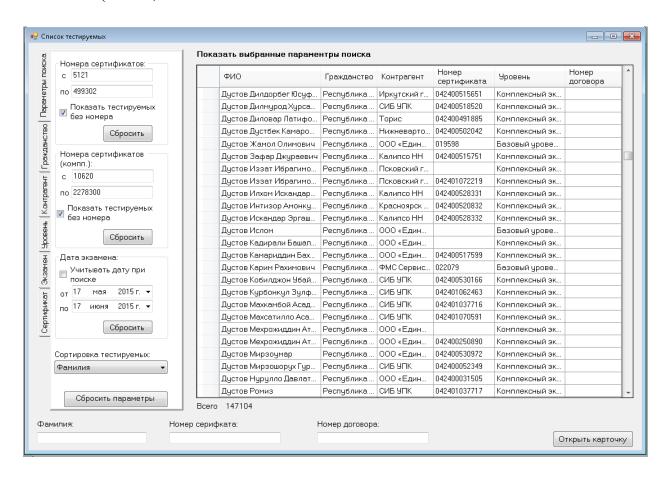
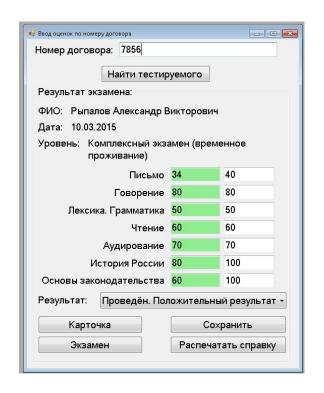


Рис. 8. Список тестируемых

Поиск по таблице осуществляется с помощью полей для ввода внизу окна: поиск по фамилии, поиск по номеру сертификата и поиск по номеру заключенного договора.

Ввод оценок за экзамен

Функция ввода (Рис. 9) оценок позволяет по номеру договора однозначно определить нужный экзамен и внести оценки. При вводе любой оценки происходит пересчет результата экзамена. Если оценка за субтест превышает необходимый минимум, поле подсвечивается зеленым, в противном случае – красным. В случае успешной сдачи экзамена сертификат становится доступным для печати.



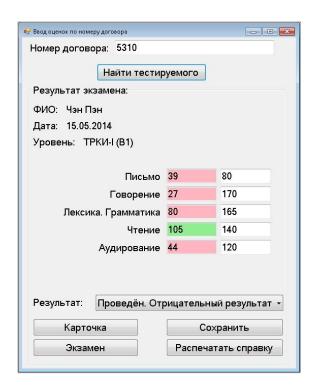


Рис. 9. Карточка ввода результатов проведенного тестирования

Загрузка данных

Данные в виде xls и csv файлов, составленных по необходимому шаблону, доступны для загрузки в базу (Рис. 10). После выбора файла и контрагента, идет проверка данных на полноту и корректность. В случае нахождения ошибок появляется предупреждение, и загрузка прекращается.

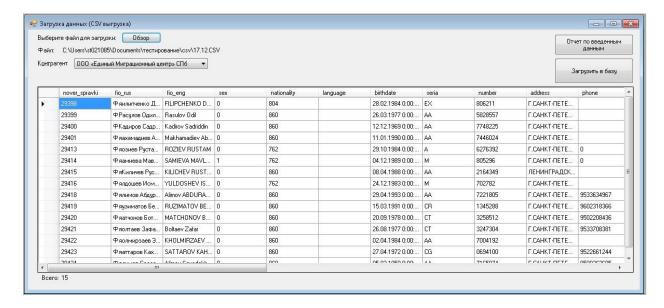


Рис. 10. Окно загрузки файла с ведомостью по проведенному тестированию в формате csv

Поскольку ошибки в данных обнаруживаются на этапе загрузки в базу, необходимо проводить такую проверку непосредственно перед передачей данных в ЦЯТ.

Печать сертификатов

К печати сертификатов (Рис. 11) допускаются сертификаты, готовые к печати и ожидающие номера. При печати каждому сертификату присваивается номер из введенного диапазона, создается протокол печати, печатаются справки, если это необходимо.

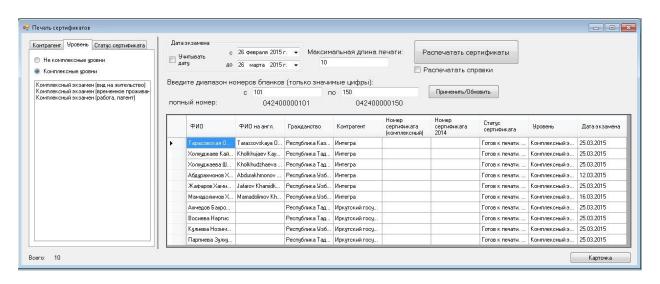


Рис. 11. Окно печати сертификатов

Протоколы

Помимо протокола, печатается реестр выдачи сертификатов.

№ сертификата	ФИО	Страна	Уровень	Дата рождения	Паспорт	Дата экзамена	Дата выдачи	Подпись выдавшего	Подпись получившего
042400500707	Кузелева Елена	Украина	Комплексный экзамен (вид на жительство)	04.01.1983	***	10.03.2015			
042400500708	Рахматов Фарход Тагоймуродович	Республика Таджикистан	Комплексный экзамен (вид на жительство)	03.07.1981	***	21.03.2015			

Протокол № 2134 от 24.03.2015

Заседания комиссии Санкт-Петербургского государственного университета по проведению государственного тестирования иностранных граждан и лиц без гражданства по русскому языку как иностранному

1. Считать следующих лиц успешно прошедшими государственное тестирование по русскому языку и выдать им сертификат «О прохождении государственного тестирования по русскому языку», удостоверяющим владение русским языком, знание истории России и основ законодательства Российской Федерации

Кузелева Елена, Украина, № 042400500707

Рахматов Фарход Тагоймуродович, Республика Таджикистан, № 042400500708

Председатель комиссии по тестированию:

Т.А. Шутова

Реестр заключенных договоров

ФИО	Номер договора	Дата договора	Дата рождения	Паспорт	Гражданство	Цена	Дата экзамена
Карамян Севак Гамлетович	988 - 11 / 8745	25.03.2015	15.03.1981	***	Республика Армения	5000	02.04.2015
Восканян Рудик Месропович	988 - 11 / 8749	25.03.2015	06.08.1985	***	Республика Армения	5000	16.04.2015
Вердеш Ион	988 - 11 / 8750	25.03.2015	23.11.1968	***	Республика Молдова	5300	28.03.2015
Меребашвили Гиоргий	988 - 11 / 8751	25.03.2015	25.12.1975	***	Грузия	5000	02.04.2015
Чой Юн Сук	988 - 11 / 8752	25.03.2015	18.07.1997	***	Республика Корея	5800	02.04.2015
Антия Гиоргий	988 - 11 / 8753	25.03.2015	29.06.1976	***	Абхазская АССР	5000	02.04.2015
Ясюкевич Николай Георгиевич	988 - 11 / 8754	25.03.2015	01.08.1978	***	Республика Беларусь	5000	09.04.2015
Джумабаева Муборак	988 - 11 / 8755	25.03.2015	17.04.1992	***	Туркменистан	5300	27.03.2015
Син Лие Куен Джослин	988 - 11 / 8757	25.03.2015	10.11.1986	***	Сингапур	5300	09.04.2015
Лю Ян	988 - 11 / 8759	25.03.2015	02.11.1995	***	Китай	5300	09.04.2015

Отчетность контрагентов

Для сверки количества тестируемых реализована функция подсчета проведенных экзаменов за указанный период для выбранного контрагента (Рис. 12).

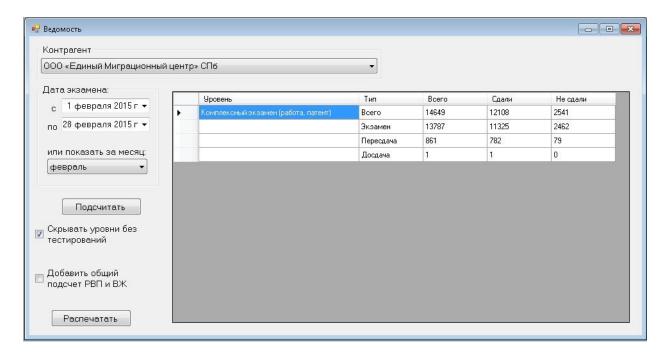


Рис. 12. Отчетность контрагентов

И вывод результата в Word:

Уровень:	Тип экзамена:	Всего	Тест успешно пройден:	Тест НЕ пройден:
	экзамена.	тестировалось:	проиден.	проиден.
Комплексный экзамен (работа, патент)	Всего	14649	12108	2541
	Экзамен	13787	11325	2462
	Пересдача	861	782	79
	Досдача	1	1	0

ООО «Единый Миграционный центр» СПб за февраль

Модуль статистического анализа и аналитические отчеты

Модуль статистического анализа

Построение графиков распределения различных показателей в соответствии с выбранными критериями доступно через пункт меню «Статистика» – «Полная статистика» (Рис. 13-15). Критерии для выбора: год тестирования, пол, гражданство И возраст тестируемого, тестирования. Подсчет статистик ведется в следующих направлениях: распределение тестируемых ПО возрастным группам, проведенных тестирований, по гражданству, по уровням тестирования и по количеству попыток сдачи экзамена.

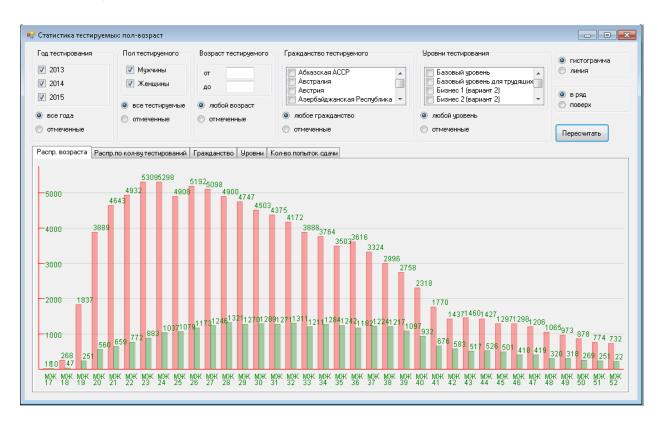


Рис. 13. Окно «Полная статистика» (закладка – «распределение по возрасту»)

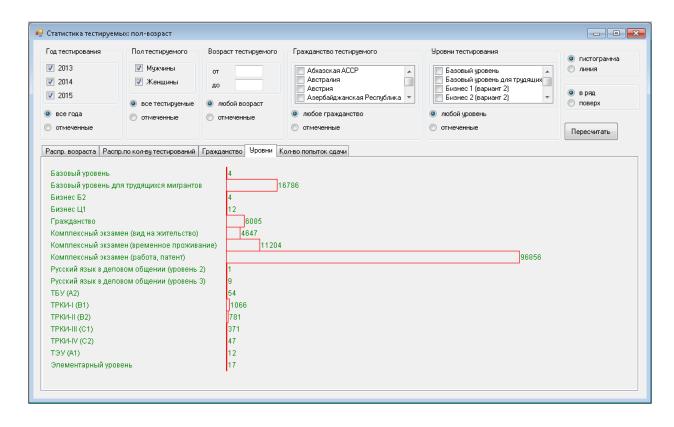


Рис. 14. Окно «Полная статистика» (закладка – «распределение по уровням тестирования»)

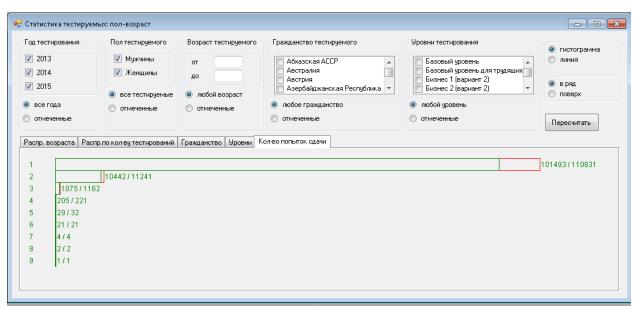


Рис. 15. Окно «Полная статистика» (закладка — «количество попыток сдачи тестирования»)

<u>Изменение количества проведенных тестирований при изменении проходных</u> баллов экзамена.

Для каждого уровня тестирования можно задать минимальные допустимые баллы, необходимые для успешной сдачи экзамена, и пересчитать количество проведенных тестирований при новых значениях. Помимо информации об изменении количества тестирований, также выводятся сведения о странах, которые в большей степени подвергаются изменениям. Пример работы функционала указан на рисунке 16.

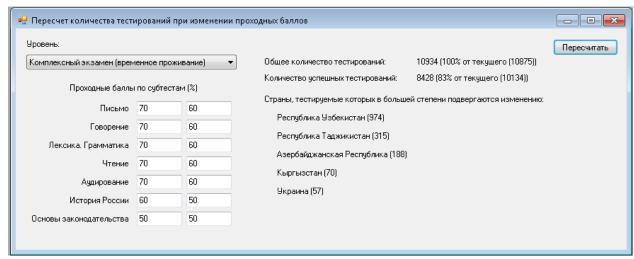


Рис. 16. Изменение количества тестирований при переопределении проходных баллов

Аналитические отчеты

Программный комплекс позволяет строить аналитические отчеты:

- Отчет по форме-1
- Отчет по форме-2
- Отчет по форме-3

Отчет по форме-1

Данный отчет отображает количество общее количество тестируемых и успешно прошедших тестирование в разрезе стран (Рис. 17). Реализована возможность вывода информации в виде Excel-файла.

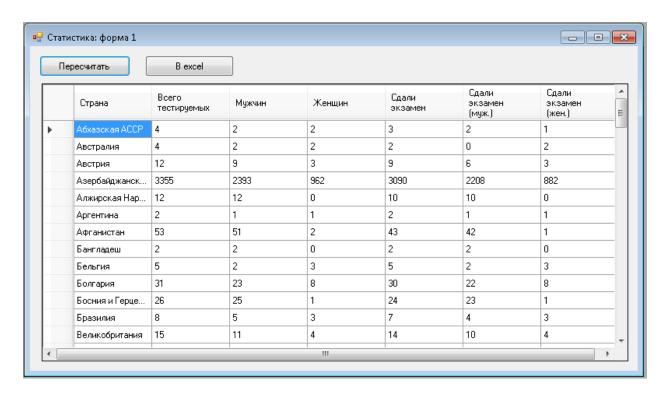


Рис. 17. Пример отчета по форме-1

Отчет по форме-2

В отчете отображена информация относительно выбранной страны: количество граждан этой страны, проходивших тестирование, с разделением на мужчин и женщин, минимальный, средний и максимальные возраста тестируемых, средние баллы и прочее (Рис. 18).

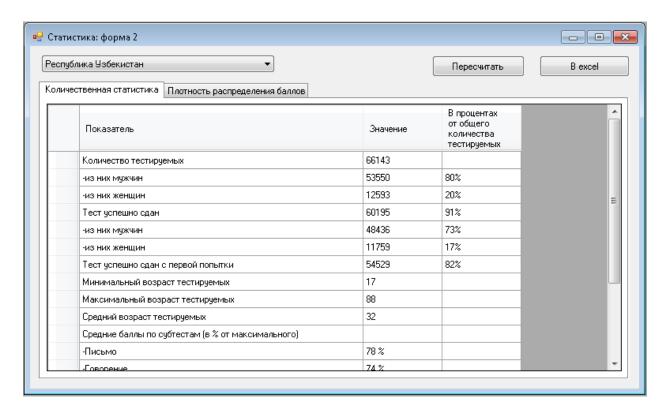


Рис. 18. Пример отчета по форме-2

Отчет также отображает плотность распределения баллов по выбранному уровню и компоненту экзамена (Рис. 19).

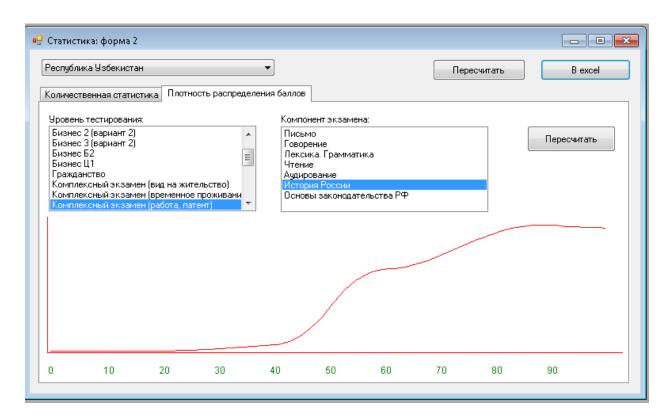


Рис. 19. Пример отчета по форме-2. Плотность распределения оценок

Отчет по форме-3

Отчет предоставляет сведения о динамике средних показателей по компонентам экзамена для выбранного уровня тестирования (Рис. 20). Возможна сортировка данных по любому параметру (количество тестируемых, общий средний показатель или средний балл по компоненту экзамена). Отчет составляется для стран, с большим общим количеством тестируемых или для стран, с большим количеством тестируемых по выбранному уровню.

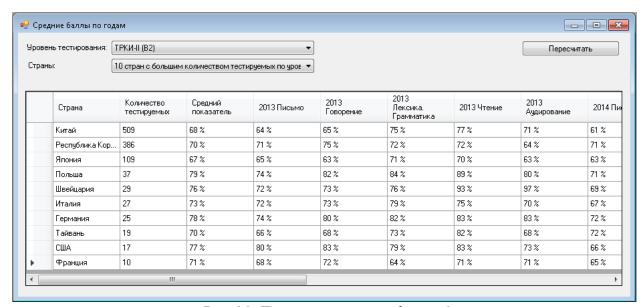


Рис. 20. Пример отчета по форме-3

3. Аналитические отчеты

Отчеты в программе строятся автоматически, в соответствии с выбранными критериями.

Отчет по форме – 1.

Отчет по форме – 1 составляется по полному перечню стран, граждане которых проходили тестирование на знание русского языка. Данный отчет отображает общее количество тестируемых и количество успешно сдавших тестирование (с разделением по полу).

Страна	· · · · · · · · · · · · · · · · · · ·	ее количесп естируемых	Успешно сдавшие тестирование				
	Всего	Мужчин	Женщин	Всего	Мужчин	Женщин	
Абхазская АССР	4	2	2	4	2	2	
Австралия	4	2	2	2	0	2	
Австрия	12	9	3	8	5	3	
Азербайджанская Республика	3286	2334	952	3021	2152	869	
Алжирская Народная Демократическая Республика	12	12	0	10	10	0	
Аргентина	2	1	1	2	1	1	
Афганистан	51	50	1	42	41	1	
Бангладеш	2	2	0	2	2	0	
Бельгия	5	2	3	5	2	3	
Болгария	31	23	8	30	22	8	
Босния и Герцеговина	26	25	1	24	23	1	
Бразилия	8	5	3	7	4	3	
Великобритания	14	10	4	13	9	4	
Венгрия	5	3	2	4	3	1	
Венесуэла	1	1	0	0	0	0	
Вьетнам	100	56	44	96	55	41	
Гана	2	2	0	2	2	0	
Германия	64	38	26	50	30	20	
Греция	19	12	7	17	11	6	
Грузия	262	147	115	230	130	100	
Дания	1	0	1	1	0	1	
Египет	33	33	0	29	29	0	

Израиль	17	16	1	13	12	1
Индия	12	11	1	11	10	1
Индонезия	4	4	0	2	2	0
Иордания	6	4	2	5	3	2
Ирак	4	4	0	4	4	0
Иран	9	6	3	7	4	3
Ирландия	8	5	3	4	3	1
Испания	22	13	9	18	9	9
Италия	78	40	38	61	35	26
Йеменская Республика	3	2	1	2	2	0
Камерун	2	2	0	2	2	0
Канада	9	6	3	6	5	1
Кипр	3	3	0	3	3	0
Китай	879	397	482	485	197	288
КНДР	80	80	0	80	80	0
Колумбия	3	2	1	2	2	0
Кот-д'Ивуар	1	1	0	1	1	0
Куба	2	1	1	2	1	1
Кыргызстан	3284	2186	1098	3079	2037	1042
Лаос	2	2	0	1	1	0
Латвия	119	74	45	115	71	44
ЛБГ	157	105	52	144	96	48
Ливан	4	4	0	3	3	0
Литва	32	16	16	28	14	14
Македония	5	4	1	5	4	1
Мали	1	1	0	1	1	0
Марокко	9	8	1	8	7	1
Мексика	2	1	1	2	1	1
Монголия	2	0	2	2	0	2
Непал	2	2	0	2	2	0
Нигерия	11	10	1	7	7	0
Нидерланды	9	5	4	5	4	1
Норвегия	2	2	0	2	2	0
Пакистан	10	10	0	9	9	0
Палестина	16	16	0	14	14	0
Перу	2	1	1	1	0	1
Польша	88	40	48	74	33	41
Португалия	1	1	0	1	1	0
Республика Абхазия	10	7	3	7	5	2
Республика Албания	1	1	0	0	0	0
Республика Армения	1743	958	785	1646	902	744
Республика Беларусь	217	97	120	208	91	117

Республика Казахстан	940	471	469	919	458	461
Республика Корея	200	92	108	111	57	54
Республика Молдова	7248	4606	2642	7022	4427	2595
Республика Таджикистан	22767	19883	2884	20379	17717	2662
Республика Узбекистан	63603	51534	12069	57932	46655	11277
Республика Эквадор	1	1	0	1	1	0
Россия	6	4	2	4	3	1
Румыния	1	1	0	1	1	0
Сент-Китс и Невис	1	0	1	1	0	1
Сербия	29	25	4	25	21	4
Сингапур	2	0	2	2	0	2
Сирийская Арабская Республика	31	31	0	28	28	0
Словакия	14	4	10	12	4	8
Словения	2	0	2	1	0	1
Судан	3	3	0	2	2	0
США	71	43	28	47	29	18
Таиланд	23	11	12	12	6	6
Тайвань	26	8	18	18	5	13
Тунис	11	10	1	8	7	1
Туркменистан	133	93	40	94	60	34
Турция	96	87	9	85	80	5
Уганда	4	4	0	4	4	0
Украина	13168	9605	3563	12941	9438	3503
Уоллис и Футуна	1	1	0	1	1	0
Филиппины	2	2	0	1	1	0
Финляндия	16	9	7	13	9	4
Франция	31	20	11	24	17	7
Хорватия	3	2	1	1	0	1
Черногория	1	1	0	1	1	0
Чехия	33	18	15	32	17	15
Швейцария	66	53	13	21	9	12
Швеция	6	4	2	6	4	2
Шри-Ланка	4	2	2	4	2	2
Эстония	69	44	25	67	44	23
Ямайка	1	0	1	0	0	0
R иноп R	141	74	67	99	54	45

Табл. 3. Пример отчета по форме – 1

Отчет по форме - 2

Отчет по форме — 2 отображает детальную информацию для каждой страны или всех стран вместе. Включает в себя следующую информацию — общее количество тестируемых (всего, отдельно мужчин и отдельно женщин), количество успешно прошедших тестирование (всего, отдельно мужчин, отдельно женщин), количество успешно сдавших тестирование с первой попытки, минимальный, максимальный и средний возраст тестируемых, среднее количество баллов по каждому компоненту экзамена.

Показатель	Значение критерия	В процентах от общего количества тестируемых
Количество тестируемых	119673	
-из них мужчин	93751	78%
-из них женщин	25922	22%
Тест успешно сдан	109683	91%
-из них мужчин	85454	71%
-из них женщин	24229	20%
Тест успешно сдан с первой попытки	99642	83%
Минимальный возраст тестируемых	16	
Максимальный возраст тестируемых	93	
Средний возраст тестируемых	32	
Средние баллы по субтестам (в % от максимального)		
-Письмо	80%	
-Говорение	78%	
-Лексика. Грамматика	81%	
-Чтение	84%	
-Аудирование	79%	
-История России	76%	
- Основы законодательства РФ	76%	

Табл.4. Пример отчета по форме-2 для всех стран

Отчет по форме-3

Отчет по форме-3 отображает средние баллы по компонентам экзамена выбранного уровня тестирования для стран с наибольшим количеством тестируемых в целом или по уровню за 2013-2015 года. Отчет дает наглядное изображение тенденций изменения или сохранения среднего уровня с течением времени.

Страна	Кол- во тест.	Ср. знач.	2013 Пс	2013 Гв	2013 Гр	2013 Чт	2013 Ауд	2014 Пс	2014 Гв	2014 Гр	2014 Чт	2014 Ауд
Республика Узбекистан	10080	80%	88%	85%	81%	92%	91%	73%	74%	81%	83%	81%
Республика Таджикистан	1804	78%	88%	87%	79%	91%	89%	71%	72%	79%	81%	78%
Украина	1402	93%	96%	99%	96%	98%	97%	87%	89%	96%	93%	91%
Республика Молдова	1297	88%	93%	97%	89%	98%	92%	81%	84%	93%	90%	87%
Кыргызстан	733	85%	88%	87%	79%	94%	90%	79%	82%	83%	88%	85%
Республика Армения	357	92%	92%	95%	91%	96%	96%	88%	91%	89%	93%	91%
Азербайджанская Республика	232	86%	89%	87%	84%	92%	93%	80%	82%	82%	87%	86%
Китай	58	90%	91%	77%	80%	95%	98%	94%	85%	80%	96%	98%

Табл.5. Пример отчета по форме — 3. Распределения средних баллов по экзамену «Базовый уровень для трудящихся мигрантов» за 2013-2014 года. Баллы за 2015 год отсутствуют, так как экзамен отменен с 01.01.2015. (Пс — письмо, Гв — говорение, Гр — Грамматика, Чт — чтение, Ауд — аудирование, Ист — история РФ, Зак — основы законодательства РФ)

Страна	Кол- во тест.	Ср. знач.	2013 Пс	2013 Гв	2013 Гр	2013 Чт	2013 Ауд	2014 Пс	2014 Гв	2014 Гр	2014 Чт	2014 Ауд	2015 Пс	2015 Гв	2015 Гр	2015 Чт	2015 Ауд
Украина	1381	93%	91%	97%	97%	95%	88%	90%	97%	95%	93%	87%	91%	96%	96%	94%	87%
Республика Узбекистан	1182	89%	88%	95%	92%	89%	82%	86%	94%	92%	90%	82%	89%	96%	93%	92%	84%
Республика Таджикистан	782	79%	77%	84%	75%	79%	74%	78%	84%	78%	82%	76%	77%	85%	79%	84%	76%
Республика Армения	765	83%	79%	87%	81%	83%	76%	79%	88%	82%	89%	80%	79%	89%	84%	88%	81%
Республика Молдова	620	88%	84%	94%	91%	90%	82%	82%	94%	91%	91%	83%	82%	94%	90%	89%	82%
Азербайджанская Республика	471	79%	76%	85%	81%	83%	74%	75%	84%	79%	84%	77%	65%	81%	72%	80%	72%
Грузия	210	84%	80%	89%	85%	88%	79%	77%	90%	85%	87%	78%	77%	91%	83%	87%	78%
ЛБГ	148	87%	82%	90%	88%	82%	74%	86%	93%	93%	88%	84%	84%	95%	91%	85%	84%
Республика Казахстан	106	93%	89%	97%	94%	93%	84%	92%	96%	96%	92%	84%	93%	98%	97%	94%	89%
Латвия	72	89%	94%	95%	100%	96%	90%	87%	94%	95%	80%	79%	91%	95%	97%	84%	87%

Табл.6. Пример отчета по форме – 3. Распределения средних баллов по экзамену «Гражданство» за 2013-2015 года. (Пс – письмо, Гв – говорение, Гр – Грамматика, Чт – чтение, Ауд – аудирование, Ист – история РФ, Зак – основы законодательства РФ)

4. Вероятностно-статистический анализ данных

<u>Анализ зависимости успешной сдачи тестирования от персональных данных</u> <u>тестируемого</u>

Всего за 2013-2015 год в центре тестирования зарегистрировано более 100 тысяч человек. Возраст тестируемых варьируется от 16-17 лет до 68 лет (Рис. 21). Большая часть тестируемых — мужчины и женщины в возрасте от 20 до 35 лет, причем мужчин практически в три раза больше, чем женщин.



Рис. 21. Распределение тестируемых в срезе по возрасту (красный – мужчины, зеленый – женщины)

Большая часть тестируемых является гражданами Узбекистана. Далее идут граждане Таджикистана, Украины, Республики Молдова, Кыргызстана, Азербайджана, Китая, Казахстана и др (Рис. 22).

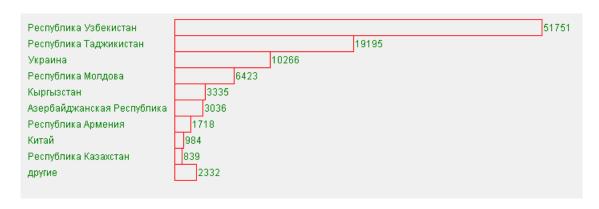


Рис. 22. Распределение тестируемых в срезе по гражданству за 2013-2015 года

До введения комплексного экзамена с 1-го января 2015 года количество тестируемых не превышало 400 человек в месяц. Однако, в январе 2015 года общее количество иностранных граждан, желающих пройти тестирование, достигло 25000 (Рис. 23). Можно заметить, что большинство тестируемых успешно сдают экзамен с первой попытки.

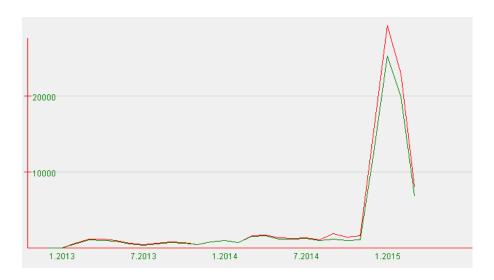


Рис. 23. Распределение тестируемых в срезе по дате проведения тестирования (красный – общее количество тестируемых, зеленый – количество успешно прошедших тестирование)

Очевидно, большая часть тестируемых сдает комплексный экзамен, а также (в 2013-2014 годах) экзамен базового уровня для мигрантов (ТБУМ) и на гражданство (Рис. 24).

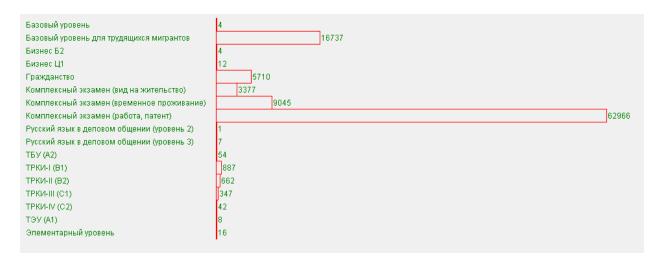


Рис. 24. Распределение тестируемых в срезе по уровню тестирования за 2013-2015 года

Практически 80% тестируемых сдают экзамен с первой попытки, еще 8% - со второй, 10% тестируемых не могут сдать экзамен успешно (Рис. 24).



Рис. 24. Распределение тестируемых в срезе по количеству попыток сдачи с указанием количества тестируемых, успешно прошедших тестировнание

Построение графиков плотности распределения выполняется в соответствии с заданными критериями: для конкретного уровня или группы уровней, по определенному набору компонент экзамена и для определенной страны (Рис. 25-26).

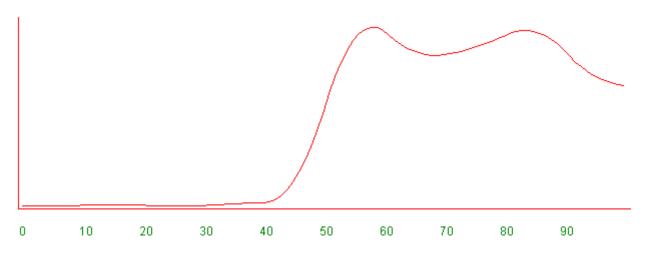


Рис. 25. Плотность распределения количества баллов (в процентом соотношении к максимальному баллу) для граждан Республики Узбекистан по компоненту «История России» комплексного экзамена (уровень – РВП, проходной балл 50%)

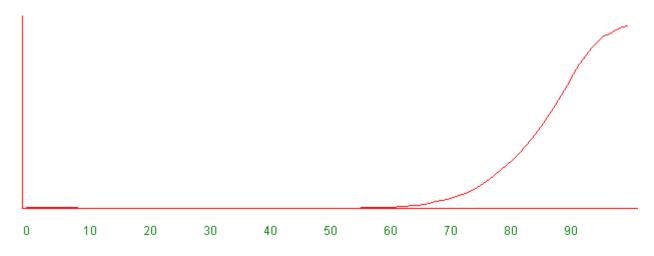


Рис. 26. Плотность распределения количества баллов (в процентном соотношении к максимальному баллу) для граждан Украины по компоненту экзамена «Письмо» комплексного экзамена (уровень – РВП, проходной балл 70%)

По представленным плотностям распределения количества баллов (рис. 25-26) можно видеть, что при увеличении проходного балла в первом случае, число успешно сдавших тестирование сократится на значительную величину. К примеру, при установлении проходного балла в 60%, т.е. его увеличении на 10%, количество успешно прошедших тестирование уменьшится на 970 человек. В то время как во втором случае – практически не изменится.

<u>Прогнозирование количества тестируемых и процента успешного</u> прохождения.

Прогнозирование временных рядов заключается в построении модели для предсказания будущих событий основываясь на известных событий прошлого, предсказания будущих данных до того как они будут измерены.

На основании данных тестирования за прошедшие 2013-2014 года (Рис. 27), можно строить предположения о количестве тестируемых на ближайший год.

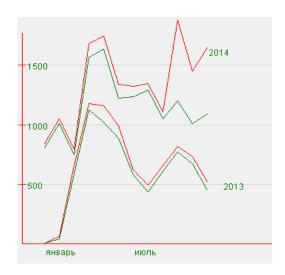


Рис. 27. Распределение тестируемых в срезе по дате проведения тестирования

С помощью модели Хольта-Винтерса был получен прогноз относительно количества тестируемых на ближайшие месяцы. Точность прогноза составила 89%.

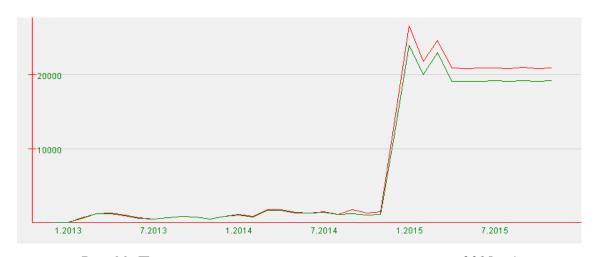


Рис. 28. Прогнозирование количества тестируемых на 2015 год

<u>Изменение количества проведенных тестирований при изменении проходных</u> баллов экзамена.

Изменяя минимальные требования, необходимые для успешной сдачи экзамена, можно проанализировать изменение количества проведенных тестирований. Уменьшая проходной балл, мы тем самым увеличиваем вероятность успешного прохождения тестирования и снижаем общее количество проводимых тестирований. Увеличивая проходной балл, наоборот — уменьшаем вероятность сдачи теста и увеличиваем общее число тестирований.

На данный момент действуют следующие проходные баллы:

Уровень тестирования	Модуль	Модуль	Модуль
	«Русский язык»	«История Р Φ »	«Основы
			законодательства»
Комплексный экз. (ИР)	60% (один компонент 50%)	50%	50%
Комплексный экз. (РВП)	70% (один компонент 60%)	50%	50%
Комплексный экз. (ВЖ)	80% (один компонент 70%)	75%	75%
Остальные уровни	66% (один компонент на 60%)	-	-

В базе данных хранятся баллы по всем компонентам экзамена для каждого проведенного тестирования, это делает возможным переопределение результата экзамена при изменении проходных баллов. Таким образом, тестируемый, не сдавший экзамен, при более низком пороге мог бы иметь «положительный» результат, а все его дальнейшие пересдачи не имели бы места, следовательно, общее количество проводимых тестирований сократилось бы.

В приложении можно задавать проходной балл по каждому компоненту экзамена по отдельности для конкретного уровня тестирования и наблюдать, как изменится количество тестирований по сравнению с текущим

значением. Например, если для комплексного экзамена (уровень — РВП) увеличить проходной балл по истории и основам законодательства на 10% (60% вместо 50%), то количество мигрантов, успешно прошедших тестирование, сократится на 25% (при 60% — 7665 положительных результатов, при текущем значении — 10115). В основном, это будут граждане Республики Узбекистан (1380 человек), Республики Таджикистан (450 человек), Азербайджанской Республики (265 человек), Кыргызстана (90 человек) и Украины (80 человек).

Однако, снизив планку на 5% от установленного значения (до 45%) мы получим сокращение общего числа тестирований всего на 1% (при 45% – 10829, при текущем значении – 10945).

Выводы

В ходе исследования результатов тестирований иностранных граждан по русскому языку были сделаны следующие основные выводы:

- Задания по каждому компоненту экзамена разработаны таким образом, что они являются доступными для сдачи иностранными гражданами. Средние показатели по всем субтестам находятся на приблизительно одинаковом уровне (80%). Однако можно заметить, что для компонентов «Истории Российской Федерации» и «Основам законодательства Российский Федерации» это значение ниже, чем соответствующие значения для других компонентов (табл. 2). Это может быть связано как с новизной самих компонентов, так и возможной недоступностью типовых заданий в свободном доступе.
- Наиболее подготовленными к экзаменам оказываются граждане иностранных государств, на территории которых распространен русский язык Украины и Республики Молдова. Средние баллы для этих стран довольно высоки 88-98%. Наименее подготовленными оказываются граждане Республики Таджикистан, их средние показатели колеблются в интервале 75-80%.
- Исследуя плотности распределения результатов экзамена, отметим, что нет ярко выраженных групп с резкими перепадами подготовленности к тестированию. Изменение количества баллов имеет сглаженный, плавный характер. Причем, для каждой страны плотности распределения баллов по разным компонентам экзамена имеют схожий вид, что говорит о равной подготовке ко всем частям тестирования.
- В ходе исследования было получено, что динамика количества тестируемых за 2013-2015 года имеет одинаковый характер изменения, обладает ярко выраженной сезонностью и трендом.

Также отметим, что коэффициент, равный отношению количества граждан, успешно сдавших тестирование, к общему количеству тестируемых мало изменяется с течением времени.

• Анализ и прогнозирование количества тестируемых за период 2013-2015 годов, проведенные с помощью трехпараметрической модели Хольта-Винтерса, показали, что даже при резком увеличении количества тестируемых (среднее количество тестируемых в месяц: в 2013г. – 760 человек, 2014 г. – 1250 человек, 2015 г.– 21400 человек) характер изменения динамики сохраняется.

Программный комплекс «База сертификатов по русскому языку» успешно используется сотрудниками Центра Языкового Тестирования с декабря 2014 года.

Естественным продолжением программы является создание программного обеспечения для проведения тестирования в электронном виде.

Заключение

В ходе исследования были решены следующие задачи:

- проведен анализ предметной области: миграционной политики и тестирования иностранных граждан по русскому языку, с выделением следующих сущностей: страны, уровни тестирования, предметы и компоненты экзамена, тестируемые.
- На основе анализа предметной области построена информационнологическая модель, которая в свою очередь, является основой для проектирования информационно-аналитического комплекса и проведения вероятностно-статистического анализа результатов тестирования.
- Спроектирован и реализован информационно-аналитический программный комплекс для проведения тестирования с использованием языков программирования С# и SQL, технологии ASP .NET и платформ ASP .NET MVC и Entity Framework.
- Разработана система составления аналитической отчетности в разрезах: страна, предмет, год тестирования, индивидуальные характеристики тестируемого: пол, возраст, гражданство и т.д.
- Разработана методология проведения вероятностно-статистического анализа результатов тестирования иностранных граждан, анализ и прогнозирование количества тестируемых и результатов их прохождения тестов.
- На основе проведенного анализа с использованием информационноаналитического программного комплекса получены результаты, которые описывают процесс тестирования в целом, и сделаны соответствующие выводы, которые могут служить поддержкой для принятия решений с целью совершенствования процесса тестирования.

Таким образом, цели и задачи, поставленные в диссертации, а именно: создание информационно-аналитического программного комплекса с рассматриваемым функционалом, разработка методологии проведения вероятностно-статистического анализа и прогнозирования временных рядов с проверкой адекватности, были полностью решены.

Литература

- 1. Официальные статистические данные ФМС России. http://www.fms.gov.ru/about/statistics/data/
- Барковский А., Барковский И. Привлечение в Россию иностранных трудовых мигрантов: необходимость селективного подхода // Общество и экономика, 2011. № 1. С. 111-128
- 3. Метелев С.Е. Международная трудовая миграция и нелегальная миграция в России: монография. М.: Закон и право, 2006. 176 с.
- 4. Выступления и публикации представителей ФМС России. http://www.fms.gov.ru/press/publications/
- 5. Приказ № 1394 от 28 октября 2014 г. «Об утверждении перечня общеобразовательных организаций, проводящих государственное тестирование по русскому языку как иностранному языку». http://минобрнауки.рф/документы/4964
- 6. Экзамен для мигрантов: первые итоги и перспективы. http://m.ria.ru/society/20150410/1057808476.html
- 7. Описание программного комплекса «Эльтерус». http://spbeta.ru/elterus
- 8. Доступ к объектам взаимодействия Office с помощью функций языка Visual C#. https://msdn.microsoft.com/ru-ru/library/dd264733.aspx
- 9. Общие сведения о System.Data.Entity. https://msdn.microsoft.com/ru-ru/library/system.data.entity(v=vs.113).aspx
- 10. Joe Mayo. LINQ Programming. McGraw-Hill Education, 2008. 385p.
- 11. Библиотека iTextSharp. http://sourceforge.net/projects/itextsharp/
- 12. Общие сведения о ASP.NET MVC. https://msdn.microsoft.com/ru-ru/library/dd381412(v=vs.108).aspx
- 13. Приказ № 1159 от 29 августа 2014 г. «Об утверждении формы, порядка проведения экзамена по русскому языку как иностранному, истории России и основам законодательства Российской Федерации,

- требований к минимальному уровню знаний, необходимых для сдачи указанного экзамена». http://минобрнауки.pd/документы/4527
- 14. Медик В.А., Токмачев М.С.. Математическая статистика в медицине: учеб. пособие. М.: Финансы и статистика, 2007. 800 с.
- 15. Орлов А.И. Прикладная статистика. М.: Издательство «Экзамен», 2004. 656 с.

Приложение

Программный код всего комплекса очень велик, и не предоставляется возможности привести его целиком. В данном приложении представлены фрагменты кода.

<u>Функционал, отображающий изменение количества тестируемых при</u> переопределении проходных баллов тестирования.

Фрагмент кода:

```
using System;
using System.Collections.Generic;
using System.ComponentModel;
using System.Data;
using System.Drawing;
using System.Linq;
using System.Text;
using System.Windows.Forms;
namespace RuslangExam
    public partial class ChangePassed : Form
        int All_count ;
        int Passed_count ;
        long level_id;
        public ChangePassed()
            InitializeComponent();
            this.Text = "Пересчет количества тестирований при изменении проходных
баллов";
            this.MdiParent = Util.MainForm;
            FillCard();
        }
        public void FillCard()
            All count = -1;
            Passed count = -1;
            ComboServ.FillCombo(cbLevel, HelpClass.GetComboListByQuery("SELECT
CONVERT(varchar(100), Id) AS Id, PublicLevelName as Name FROM dbo.LevelName order by 2"),
false, false);
            FillExamenNames();
        public void FillExamenNames()
            int? LevelId = ComboServ.GetComboIdInt(cbLevel);
            if (LevelId.HasValue)
                using (RuslangExamEntities context = new RuslangExamEntities())
                {
                    var ExamenNames = (from x in context.ResultTarget
```

```
join n in context.ResultName on x.ResultName id
                                         equals n.Id
                                        join 1 in context. Level on x. ParentId equals 1. Id
                                          join ln in context.LevelName on l.RealLevel
                                          equals ln.Id
                                        where ln.Id == LevelId.Value
                                        select new
                                        {
                                             x.ResultName id,
                                             x.FirstMinValue,
                                             x.SecondMinValue,
                                             n.Name,
                                        }).Distinct().ToList();
                     Label[] LabelNames = { lbName1, lbName2, lbName3, lbName4, lbName5,
                    lbName6, lbName7 };
                    TextBox[] Tb1 = { tb1, tb2, tb3, tb4, tb5, tb6, tb7 };
                    TextBox[] Tb11 = { tb11, tb12, tb13, tb14, tb15, tb16, tb17 };
                    foreach (Label 1 in LabelNames)
                        1.Visible = false;
                    foreach (TextBox t in Tb1.Union(Tb11))
                        t.Visible = false;
                    foreach (var x in ExamenNames.Select(x => x.ResultName_id))
                        LabelNames[x.Value - 1].Text = ExamenNames.Where(e =>
e.ResultName_id == x).Select(e => e.Name).First();
                        var temp = ExamenNames.Where(e => e.ResultName_id == x).Select(e
=> e.FirstMinValue).First();
                        Tb1[x.Value - 1].Text = temp.HasValue? temp.Value>0 ?
temp.Value.ToString() : "66" : "66";
                        temp = ExamenNames.Where(e => e.ResultName id == x).Select(e =>
e.SecondMinValue).First();
                        Tb11[x.Value - 1].Text = temp.HasValue ? temp.Value > 0 ?
temp.Value.ToString() : "60" : "60";
                        Tb1[x.Value - 1].Visible = true;
                        Tb11[x.Value - 1].Visible = true;
                        LabelNames[x.Value - 1].Visible = true;
                    }
                }
            }
        }
        private void cbLevel_SelectedIndexChanged(object sender, EventArgs e)
            FillExamenNames();
        }
        private void btnFillGrid_Click(object sender, EventArgs e)
            long? LevelId = ComboServ.GetComboIdInt(cbLevel);
            if (All_count < 0 || Passed_count < 0 || level_id != LevelId)</pre>
            {
                All count = -1;
                Passed count = -1;
                FillInit();
            FillInit();
        }
```

Функционал создания отчета по форме-1

```
using System;
using System.Collections.Generic;
using System.ComponentModel;
using System.Data;
using System.Drawing;
using System.Linq;
using System.Text;
using System.Windows.Forms;
namespace RuslangExam
{
   public partial class Form1 : Form
        public Form1()
        {
            InitializeComponent();
            this.MdiParent = Util.MainForm;
            this.Text = "Статистика: форма 1";
        }
        public void FillGrid()
            using (RuslangExamEntities context = new RuslangExamEntities())
                NewWatch nw = new NewWatch();
                nw.SetText("Please wait");
                var N = (from x in context.Nationality
                         where x.Id != 55
                         select x).OrderBy(x=>x.Name).ToList();
                nw.SetMax(N.Count);
                nw.Show();
                DataTable tbl = new DataTable();
                DataColumn clm = new DataColumn();
                clm.ColumnName = "Страна";
                tbl.Columns.Add(clm);
                clm = new DataColumn();
                clm.ColumnName = "Всего тестируемых";
                tbl.Columns.Add(clm);
                clm = new DataColumn();
                clm.ColumnName = "Мужчин";
                tbl.Columns.Add(clm);
                clm = new DataColumn();
                clm.ColumnName = "Женщин";
                tbl.Columns.Add(clm);
                clm = new DataColumn();
                clm.ColumnName = "Сдали экзамен";
                tbl.Columns.Add(clm);
                clm = new DataColumn();
                clm.ColumnName = "Сдали экзамен (муж.)";
                tbl.Columns.Add(clm);
                clm = new DataColumn();
                clm.ColumnName = "Сдали экзамен (жен.)";
                tbl.Columns.Add(clm);
```

```
foreach (var n in N)
                {
                    nw.PerformStep();
                    DataRow rw = tbl.NewRow();
                    rw[0] = n.Name;
                    var M = (from x in context.Passport
                             where x.nationality_id == n.Id
                             && x.sex_id == 1
                             select x).Count();
                    rw[2] = M.ToString();
                    var F = (from x in context.Passport
                             where x.nationality_id == n.Id
                             && x.sex_id == 2
                             select x).Count();
                    rw[3] = F.ToString();
                    rw[1] = (M+F).ToString();
                    var TestPassd = (from x in context.Passport
                                     join s in context.Sertificate on x.ParentId equals
s.ParentId
                                     where x.nationality_id == n.Id
                                     && x.sex_id != 3
                                     && s.Sertificate_number.HasValue
                                     select x).Distinct().ToList();
                    rw[4] = TestPassd.Count.ToString();
                    rw[5] = TestPassd.Where(x=>x.sex_id ==
1).Select(x=>x).ToList().Count.ToString();
                    rw[6] = TestPassd.Where(x => x.sex_id == 2).Select(x =>
x).ToList().Count.ToString();
                    tbl.Rows.Add(rw);
                }
                for (int i = tbl.Rows.Count-1; i>=0; i--)
                    DataRow rw = tbl.Rows[i];
                    if (rw[1].ToString() == "0")
                        rw.Delete();
                dgv.DataSource = tbl;
                nw.Close();
            }
        }
        private void btnfillgrid_Click(object sender, EventArgs e)
            FillGrid();
        }
   }
}
```

Функционал вывода списка тестируемых

Фрагмент SQL –запроса к базе данных:

```
select distinct Person.Id
, Passport.FIO as 'ФИО'
, Nationality.Name as 'Гражданство'
, ContrAgent.Name as 'Контрагент'
, Sertificate.Sertificate prefix
, Sertificate.Sertificate number
, Sertificate.Sertificate_number length
, '' as 'Номер сертификата'
, LevelName.Id as 'LevelNameId'
, LevelName.PublicLevelName as 'Уровень'
, Contract.ContractNumber as 'Номер договора'
, (case when (LevelName.PublicLevelName like '%комплек%') then '1' else '0' end) as 'isComplex'
from dbo.Person
inner join dbo.Passport on Passport.ParentId = Person.Id
inner join dbo.Nationality on Passport.nationality_id = Nationality.Id
left join dbo.Sertificate on Sertificate.ParentId = Person.Id
left join dbo.Examen on Examen.ParentId = Sertificate.Id
left join dbo.ContrAgent on Sertificate.ContrAgent = ContrAgent.Id
left join dbo.Level on Sertificate.Level_id = Level.Id
left join dbo.LevelName on LevelName.Id = Level.RealLevel
left join dbo.Contract on Contract.ParentId = Examen.Id
```

Условия выборки данных составляются динамически с учетом выбранных параметров.

Функционал вывода списка заключенных договоров

Фрагмент SQL-запроса к базе данных

```
Contract.Id as ContractId
, Person.Id as PersonId
, Passport.FIO as 'ФИО'
, Contract.Date as 'Дата'
, Contract.Price as 'Цена'
, Contract.ContractNumber as 'Номер договора'
from dbo.Contract
join dbo.Examen on Contract.ParentId = Examen.Id
join dbo.Sertificate on Examen.ParentId = Sertificate.Id
join dbo.Person on Sertificate.ParentId = Person.Id
join dbo.Passport on Person.Id = Passport.ParentId
where Contract.Date>=@DateBegin
and Contract.Date<@DateEnd
and ContractNumber is not null
and Examen.ExamenStatus_id != 8
order by Contract.Date, ContractNumber
```