

ОПТИМИЗАЦИЯ ЗАЧИСЛЕНИЯ В УЧРЕЖДЕНИЯ ОБРАЗОВАНИЯ ПРИ РАБОТЕ С БОЛЬШИМИ ДАННЫМИ НА ПРИМЕРЕ АВТОМАТИЗИРОВАННОЙ СИСТЕМЫ ПОДАЧИ ЗАЯВЛЕНИЙ И ЗАЧИСЛЕНИЯ БГУИР

Владыцев В.Д., магистрант, Могилевец Д.Э., студент 353505

*Белорусский государственный университет информатики и радиоэлектроники,
г. Минск, Республика Беларусь*

Казак Т.В. – д-р психол. наук, профессор

Аннотация. Эффективный и точный алгоритм подачи заявлений и зачисления абитуриентов является одним из важнейших факторов для проведения успешной вступительной кампании. Целью работы являлась разработка эффективного алгоритм формирования списков абитуриентов, зачисленных на специальности, в зависимости от их желаний и имеющихся у них преимуществ.

Ключевые слова: АСПЗиЗ, обработка больших данных, анализ эффективности, алгоритм Гэйла-Шепли.

Введение. В современном информационном обществе, охваченном вихрём технологических инноваций, процесс зачисления абитуриентов в вузы становится всё более сложным и объёмным. Все больше образовательных учреждений стремятся использовать передовые методы обработки данных для оптимизации этого процесса. Одним из главных вызовов в сфере автоматизации процесса поступления и зачисления является эффективное управление большими данными, которые основываются на информации о кандидатах, при котором обязано обеспечиваться честность и прозрачность в принятии решений при зачислении абитуриентов.

Рассмотрим автоматизированную систему подачи заявлений и зачисления БГУИР в соответствии с правилами и порядком приёма в учреждения высшего образования утверждёнными в 2025 году. Каждый из факультетов БГУИР подразделяется на несколько специальностей. Для каждой из специальностей определяется количество бюджетных и платных мест обучения для студентов в следующем учебном году. Данное количество определяется обеспеченностью университета кадрами, количеством аудиторий, имеющимися денежными средствами, распоряжениями министерства образования и т. п. [1].

Анализ существующего алгоритма зачисления. Каждый из абитуриентов, желающих поступить в БГУИР предоставляет перечень необходимых документов в приёмную комиссию, среди них наиболее интересными с точки зрения исследования: заявление на имя руководителя УВО по установленной Министерством образования форме, оригинал аттестата об общем среднем образовании, либо оригиналы диплома о профессионально-техническом образовании и приложения к нему, либо оригиналы диплома о среднем специальном образовании и приложения к нему; оригиналы сертификатов централизованного тестирования; оригиналы сертификатов централизованного экзамена; характеристика, выданная учреждением общего среднего образования; документы, подтверждающие право абитуриента на льготы. В заявлении абитуриента помимо прочего содержится ранжированный список желаемых специальностей [2]. На данный момент имеет место следующая процедура зачисления студентов:

Первыми без вступительных испытаний в порядке перечисления зачисляются абитуриенты согласно 23 пункту Правил приема (категория 1). Затем на места, оставшиеся после зачисления абитуриентов на основании 23 пункта Правил приема, вне конкурса при наличии в документе об образовании отметок не ниже 6 (шести) баллов по предметам вступительных испытаний зачисляются абитуриенты, поступающие на основании льгот, перечисленных в 26 пункте Правил приема (категория 2). На места, оставшиеся после зачисления абитуриентов на основании 23 и 26 пунктов Правил приема, абитуриенты зачисляются по конкурсу на основе общей суммы баллов (категория 3).

Зачисление абитуриентов, участвующих в конкурсе для получения высшего образования по группе специальностей с использованием автоматизированной системы подачи заявлений и зачисления, осуществляется согласно конкурсному списку абитуриентов, сформированному в пределах группы специальностей в порядке убывания набранной абитуриентами общей суммы баллов с последующим учетом порядкового номера факультета и специальности в перечне, указанном абитуриентом в заявлении.

Автоматизированная система подачи заявлений и зачисления начинает зачисление третьей категории с абитуриента с максимальным количеством набранных баллов. Проверяет наличие

свободных мест для зачисления по первой записи в заявлении данного абитуриента. В случае если место есть – зачисляет на соответствующую специальность и факультет, иначе – переходит к следующей. И так последовательно проверяются все записи в заявлении абитуриента до тех пор, пока абитуриент не будет зачислен или не закончатся записи в его заявлении.

Если список записей исчерпан, а абитуриент не зачислен, то он попадает в промежуточный список не зачисленных. После обработки заявлений всех проранжированных абитуриентов может возникнуть две ситуации: первая – все специальности заполнены, вторая - остались вакантные места на некоторых из них. В первом случае абитуриенты из списка не зачисленных считаются не прошедшими конкурсный отбор. Во втором случае абитуриенты из списка не зачисленных снова ранжируются по сумме набранных баллов и с большими баллами зачисляются на вакантные места незаполненных специальностей решением приемной комиссии [3].

Наибольший интерес с точки зрения алгоритмической оценки представляет третья категория. Согласно приведённому выше тексту, существующий сейчас алгоритм зачисления можно представить в виде следующей блок-схемы (рисунок 1):

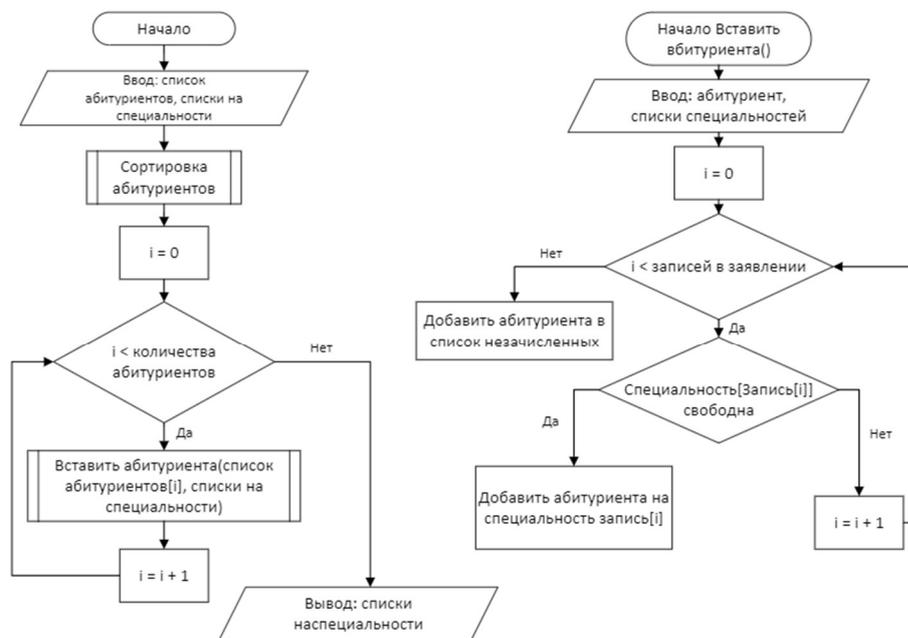


Рисунок 1 – Схема существующего алгоритма зачисления

Обозначая количество специальностей за S , количество абитуриентов, подавших заявление за A , а сложность сравнения двух абитуриентов друг с другом за C имеем. Сложность сортировки всех абитуриентов – $O(CA \log(A))$ ($O(C)$ – сложность одного сравнения, $O(A \log(A))$ – необходимое число сравнений), сложность функции вставки абитуриента в среднем – $O(S)$ (среднее количество записей в заявлении, при условии равновероятного выбора любого количества, это $\frac{S}{2}$). Тогда общая сложность алгоритма – $O(SA + CA \log(A))$.

Квадратичная сложность алгоритма, отсутствие инкрементальной обработки, ресурсоёмкость операций, латентность принятия решений, ограниченная масштабируемость и отсутствие распределённой обработки не позволяют масштабировать АСПЗиЗ на другие ВУЗы в качестве общереспубликанской системы подачи заявления и зачисления абитуриентов.

Также заметим, что данный алгоритм в представленной форме не работает “в онлайн”, то есть при добавлении новых абитуриентов в количестве ΔA необходимо поводить алгоритм с начала полностью и сложность добавления составит $O(S(A + \Delta A) + C(A + \Delta A) \log(A + \Delta A))$. Также, стоит отметить, что в приведённом выше описании не указано ничего о спорных ситуациях и о способах сообщения о них системой. Также, в нём ничего не указано о различиях в правилах зачисления для отдельных специальностей (в смысле льгот и поступления без вступительных испытаний).

Изначальный алгоритм не подходит для Big Data из-за высокой сложности, отсутствия онлайн-обработки, негибкости в отношении правил и спорных ситуаций, а также низкой масштабируемости [4]. Для эффективной работы с Big Data требуется алгоритм с линейной или близкой к линейной сложности, способностью к инкрементальной обработке и гибким правилам. Для решения проблемы был разработан алгоритм, основанный на модификации алгоритма Гэйла-Шепли.

Анализ предлагаемого алгоритма. Рассмотрим следующий возможный вариант алгоритма зачисления, который был внедрён в АСПЗиЗ БГУИР в 2024 году:

1. составить произвольный список абитуриентов.
2. первый из ещё не зачисленных абитуриентов идёт на первую желаемую специальность.
3. если он “лучше” последнего человека в списке на специальность, то он сравнивается со следующим.
4. повторяем шаг 3, пока абитуриент не окажется “хуже”, чем некоторый из абитуриентов в списке на специальность. отметим рассматриваемого абитуриента “зачисленным”
5. если длина списка больше квоты на специальность, то рассмотрим последних абитуриентов в списке, если среди них есть равные, выведем сообщение о необходимости выбора, кого оставить. иначе просто уберём из списка последнего, отметив, что он не может поступить на эту специальность.
6. убранный абитуриент отметим не зачисленным.
7. если абитуриент “хуже” последнего человека на первой желаемой специальности, то перейдём к следующей желаемой.
8. если желаемые специальности закончились, отметим рассматриваемого абитуриента как “не поступившего”
9. перейдём к следующему абитуриенту, если он есть.
10. повторим процесс сначала для не зачисленных абитуриентов, если они есть.

В данном контексте слова “лучше” и “хуже” означают сравнение в смысле набранных на экзаменах баллов, льгот, среднего балла аттестата и т.п.

Для оптимизации работы и ускорения процесса возможен отдельный запуск алгоритма по отфильтрованным группам «Вне конкурса», «Без экзаменов» и «По конкурсу». Процесс можно представить в виде следующей блок схемы (рисунок 2).

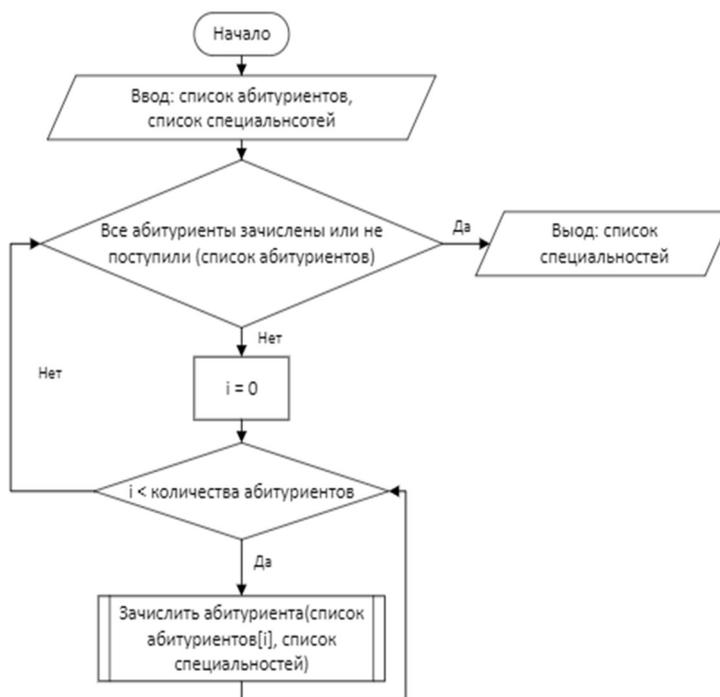


Рисунок 2 – Блок-схема предлагаемого алгоритма

Обозначая максимальную квоту на специальность среди всех за M имеем. Средняя сложность функции зачисления нового абитуриента составляет $O(S + CM)$, поскольку, в сущности, она проходит по всем желаемым абитуриентом специальностям и для одной и только одной делает M действий сравнения и M действий для вставки абитуриента. Поскольку $S \ll M$ (количество специальностей в университете значительно меньше, чем средняя максимальная квота), то конечная сложность алгоритма зачисления одного абитуриента: $O(CM)$. Для упрощения понимания работы функцию следует рассмотреть следующую блок схему (рисунок 3).

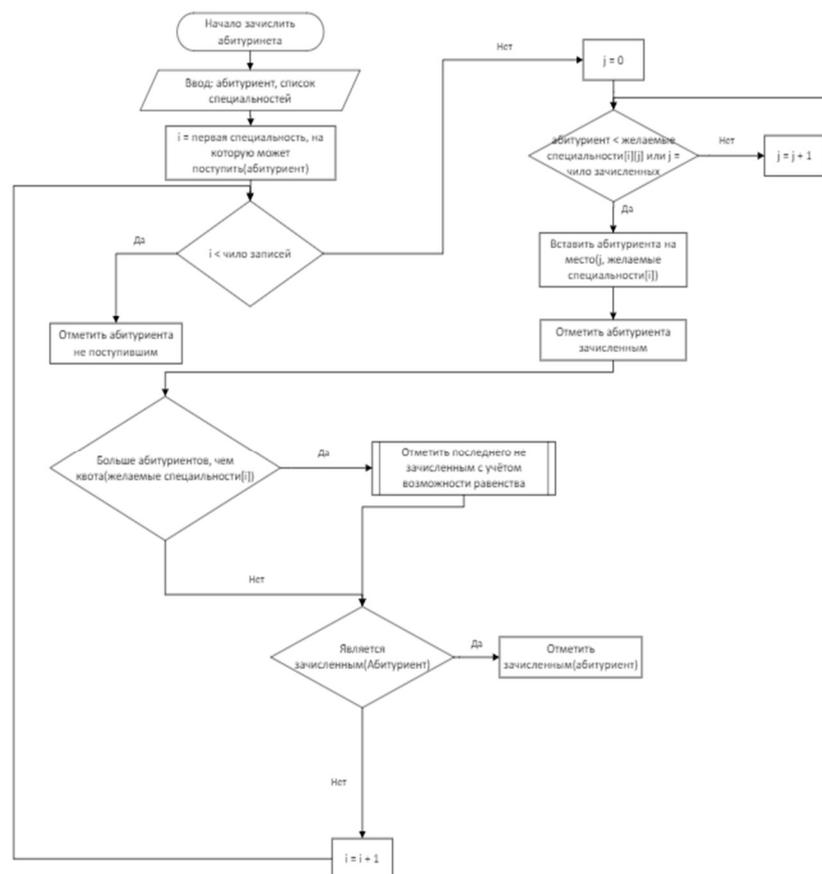


Рисунок 3 – Блок-схема функции, используемой в предлагаемом алгоритме

Сложность внутреннего цикла основного алгоритма составляет $O(ASM)$, поскольку для каждого из A абитуриентов функция его зачисления вызывается всего один раз. Заметим, что потребуется не более чем S проходов цикла, чтобы все абитуриенты были помечены либо как зачисленные, либо как не поступившие. Тогда сложность алгоритма зачисления – $O(ASCM)$. Преимуществами данного алгоритма являются линейная сложность от числа абитуриентов и возможность эффективной работы “в онлайн”. Действительно, если заявление подали ещё ΔA абитуриентов, то достаточно просто провести процедуру алгоритма с имеющимися списками зачисленных. Тогда сложность добавления группы абитуриентов размером ΔA составит $O(\Delta ASCM)$. Также, в алгоритме учтена возможность равенства нескольких абитуриентов. Также, стоит отметить, что алгоритм даёт возможность учесть различия в правилах приёма на разных специальностях, для этого достаточно просто изменить функцию сравнения двух абитуриентов для конкретной специальности и использовать её при попытке зачислить нового абитуриента на данную специальность.

Важным недостатком стоит считать большое значение константы от числа абитуриентов – SCM .

Можно заметить некоторую аналогию с алгоритмом Гэйла-Шепли формирования устойчивого бракосочетания. Действительно, если рассматривать абитуриентов как “мужей”, а специальности как “жён”, то приведённый алгоритм действительно соответствует алгоритму Гейла-Шепли с добавления возможности “полиандрии”. Как известно, алгоритм Гэйла-Шепли гарантирует наилучшее устойчивое бракосочетание для женихов [5], что вполне согласуется с рациональным требованием того, что каждый из абитуриентов должен поступить на лучшую из специальностей, на которые может и хочет поступить. Т.е., более формально, если некоторый абитуриент X не поступил на все специальности, идущие ранее специальности A в его заявлении и некоторый абитуриент Y , оказавшийся “хуже” по результатам сравнения X и Y , поступил на специальность A , то и абитуриент X поступит на специальность A .

Как можно видеть из представленных выше блок-схем, алгоритм зачисления может быть усовершенствован или переделан полностью.

Заключение. В данной работе был рассмотрен существующий алгоритм зачисления абитуриентов в БГУИР. Была оценена его алгоритмическая сложность, возможность работы “в онлайн”, отношение к спорным ситуациям. Также был предложен и рассмотрен другой, более эффективный алгоритм зачисления абитуриентов, основанный на алгоритме Гэйла-Шепли. Предложенный усовершенствованный алгоритм не является сложным для пользователей автоматизированной системы подачи заявления и зачисления, а, следовательно, сохраняет принципы прозрачности при зачислении

абитуриентов в учреждения образования, используя в разы меньше ресурсов при работе и может быть быстро внедрён в автоматизированную систему подачи заявлений и зачисления.

Список использованных источников:

1. Автоматизированная система подачи заявлений и зачисления в рамках корпорации учреждений образования / М. П. Батура [и др.] // *Высшее техническое образование: проблемы и пути развития : материалы VII Международной научно-методической конференции*, Минск, 20-21 ноября 2014 г. / Белорус. гос. ун-т информатики и радиоэлектроники ; редкол.: Е. Н. Живицкая [и др.] – Минск, 2014. – С. 18 - 19.

2. Суровцев, А. И. Исследование пользовательского опыта при взаимодействии абитуриентов с электронным кабинетом автоматизированной системы подачи заявления и зачисления БГУИР = *The study of user experience in the interaction of applicants with the electronic cabinet of the automated application and enrollment system of the BSUIR* / А. И. Суровцев, В. Д. Владымцев // *Компьютерные системы и сети : сборник статей 60-й научной конференции аспирантов, магистрантов и студентов*, Минск, 22–26 апреля 2024 г. / Белорусский государственный университет информатики и радиоэлектроники. – Минск, 2024. – С. 524–527.

3. Развитие единого информационного пространства приемной кампании Белорусского государственного университета информатики и радиоэлектроники / Никульшин Б. В. [и др.] // *Дистанционное обучение – образовательная среда XXI века: материалы XI Международной научно-методической конференции*, Минск, 12-13 декабря 2019 г. / Белорус. гос. ун-т информатики и радиоэлектроники ; редкол.: В. А. Прытков [и др.]. – Минск, 2019. – С. 219.

4. Альховик, Д. И. Big Data: проблемы и технологии / Д. И. Альховик, В. Д. Владымцев, А. Н. Марков // *BIG DATA и анализ высокого уровня = BIG DATA and Advanced Analytics : сборник научных статей IX Международной научно-практической конференции*, Минск, 17–18 мая 2023 г. : в 2 ч. Ч. 2 / Белорусский государственный университет информатики и радиоэлектроники ; редкол.: В. А. Богуш [и др.]. – Минск, 2023. – С. 201-204.

5. Гейл, Д. Поступление в колледж и стабильность брака / Д. Гейл, Л. Шеплу // *The American Mathematical Monthly*. – 1962. – Т. 69, № 1. – С. 9-15.

OPTIMIZATION OF ENROLLMENT IN EDUCATIONAL INSTITUTIONS WHEN WORKING WITH BIG DATA USING THE EXAMPLE OF THE AUTOMATED APPLICATION AND ENROLLMENT SYSTEM OF BSUIR

Vladymtsev V.D., Mahiliavets D.E.

Belarusian State University of Informatics and Radioelectronics, Minsk, Republic of Belarus

Kazak T.V. – Dr. of Sci. (Psy.), Professor

Abstract. An effective and accurate algorithm for applying and enrolling applicants is one of the most important factors for a successful admission campaign. The aim of the work was to develop an effective algorithm for the formation of lists of applicants enrolled in the specialty, depending on their desires and the advantages they have.

Keyword: Automatization, big data processing, efficiency analysis, Gale-Shapley algorithm.