

СИСТЕМНЫЙ АНАЛИЗ ПРИВЛЕКАТЕЛЬНОСТИ БАНКОВ ДЛЯ ПРЕДСТАВИТЕЛЕЙ ВИП-СЕКМЕНТА ПОТРЕБИТЕЛЕЙ В РАМКАХ ЦИФРОВОЙ ИНФОРМАЦИОННОЙ СРЕДЫ

© 2021 **Ершова Алёна Викторовна**

аспирант кафедры маркетинга

Санкт-Петербургский государственный экономический университет, Россия, Санкт-Петербург

E-mail: alena.tsyura@mail.ru

© 2021 **Родионов Дмитрий Григорьевич**

доктор экономических наук, профессор, Высшая инженерно-экономическая школа

Санкт-Петербургский политехнический университет Петра Великого, Россия, Санкт-Петербург

E-mail: dmitry.rodionov@spbstu.ru

© 2021 **Конников Евгений Александрович**

кандидат экономических наук, доцент, Высшая инженерно-экономическая школа

Санкт-Петербургский политехнический университет Петра Великого, Россия, Санкт-Петербург

E-mail: konnikov.evgeniy@gmail.com

© 2021 **Конникова Ольга Анатольевна**

кандидат экономических наук, доцент, кафедра маркетинга

Санкт-Петербургский государственный экономический университет, Россия, Санкт-Петербург

E-mail: olga.a.konnikova@gmail.com

Одним из наиболее трансформировавшихся в результате цифровизации секторов экономики стал банковский сектор. Следствием данного процесса стал значительный переход банковских услуг в онлайн. Данный переход спровоцировал трансформацию восприятия банка представителями практически всех сегментов потребителей. При этом одним из наиболее значимых в контексте данной трансформации для банка является ВИП-сегмент потребителей. Целью данного исследования является формирование актуального инструмента системного анализа банков с точки зрения их привлекательности для представителей ВИП-сегмента потребителей банковских услуг. Полученный инструмент полностью автоматизирован средствами Python, и может использоваться как банками в рамках конкурентного анализа, так и представителями ВИП-сегмента потребителей в рамках сравнительного анализа банковских организаций.

Ключевые слова: системный анализ, банк, ВИП-сегмент потребителей, Python.

Стремительный процесс цифровизации экономики является крайне дифференцированным, оказывая влияние практически на все направления потребления и производства товаров и услуг. Пандемия COVID-19 в значительной мере медицировала данный процесс, следствием чего стала трансформация множества бизнес-процессов предприятий как B2B, так и B2C направлений. Одним из секторов экономики, претерпевшим значительные технологические трансформации в рамках процесса цифровизации, стал банковский сектор. Данные трансформации затронули как внутренние бизнес-процессы, так и непосредственно процессы взаимодействия с потребителями. Значительное увеличение доли интернет-платежей, в общем объеме электрон-

ных платежей, развитие мультиинструментальных банковских приложений, предоставляющих возможность пользователям осуществлять инвестиционные операции и потребление товаров и услуг небанковского характера, а также повышение скорости и безопасности электронных платежей вследствие развития технологии блокчейн, привели к многомерной трансформации потребительских паттернов и измерения аватаров основных потребителей банковских услуг. Данные изменения оказали влияние практически на все группы потребителей, включая крайне консервативную и значимую для банковского сектора группу — потребителей ВИП-сегмента.

ВИП-сегмент потребителей банковских услуг в значительной мере отличается от иных

сегментов. Аватар представителей данного сегмента был подробно проанализирован авторами ранее, результаты данного анализа представлены в статье [1]. В первую очередь необходимо отметить следующие свойства представителей данного сегмента:

1. Центрирование надежности взаимодействия как ключевой характеристики банковской организации. В данном случае под надежностью взаимодействия можно понимать способность банка сохранять во времени в установленных пределах значения всех параметров, характеризующих оказываемые услуги [2]. Данное свойство обусловлено в первую очередь системностью потребностей представителей данного сегмента. Неспособность банка своевременно и в полном объеме оказывать требуемые услуги представителю данного сегмента потенциально провоцирует цепную реакцию, следствием которой могут стать значительные финансовые и репутационные потери, значительно превышающие альтернативные потери у представителей иных сегментов. Данное свойство во многом обеспечивается масштабируемостью и кастомизируемостью услуг, оказываемых представителям данного сегмента. Однако, вследствие цифровой трансформации, и переходе множества оказываемых банком услуг в онлайн, возможность кастомизации значительно снизилась, как и ориентация представителей ВИП-сегмента на личное взаимодействие, в рамках процесса потребления.

2. Совместимость восприятия. Данное свойство отражает минимизацию противоречий между медиа-восприятием банка основными потребителями, а также самих потребителей услуг данного банка, медиа-восприятием как представителя ВИП-сегмента, так и его возможного бизнеса и/или окружения. Данное свойство является крайне дискуссионным с точки зрения оценки, однако одними из наиболее возможных параметров квантификации данного свойства являются тональная и лексическая корреляция. Результат квантификации может базироваться на массивах естественной информации, генерируемой как банком, так и потребителями его услуг в цифровой среде.

3. Инновационная активность. Значительным следствием процесса цифровизации, для представителей практически всех сегментов потребителей банковских услуг, стало рефлексия динамики технологического развития. Данное

следствие сформировало уникальные категории ценностей, такие как относительная инновационность оказываемых услуг и динамика развития относительной Инновационности оказываемых услуг. Представители ВИП-сегмента не являются исключением, и данные свойства также являются для них ценностями.

Совокупность приведённых свойств уникально для представителей ВИП-сегмента потребителей банковских услуг. Следовательно, направленное маркетинговое воздействие на представителей данного сегмента может основываться на формировании образа банка в рамках выделенных свойств. Однако, необходимо отметить, что в следствии цифровизации, потребитель с большей вероятностью воспринимает и как следствие оценивает не сам банк, а его образ, сформированный в цифровой информационной среде. Следовательно, для целей квантификации каждого из свойств необходимо использовать исключительно открытые данные цифровой информационной среды, в первую очередь представленные в форме естественной информации. Рассмотрим последовательно квантификацию каждого из выделенных свойств.

В первую очередь рассмотрим надежность взаимодействия с банком. Данное свойство с точки зрения психологии носит реверсивный характер. Информационный фон, намеренно демонстрирующий надёжность взаимодействия как основное преимущество банка может спровоцировать сомнения в сознании представителей ВИП-сегмента. Следовательно, банк должен выделять не наличием данной информационной компоненты в общем информационном фоне, а отсутствием обратной информационной компоненты. Данная обратная информационная компонента определяется наличием в информационных единиц, определяющих технические, экономические или иные сбои в работе банка. Данная информация, вне зависимости от условной частоты появления в общем информационном фоне, неизменно формирует у представителей ВИП-сегмента бинарное восприятие банка как условно ненадежного, вследствие чего, ни один банк не будет целенаправленно генерировать данную информацию, и будет стремиться к ее максимальному сокрытию. Следовательно, источником данной информации не могут выступать официальные цифровые информационные ресурсы банка. Одним из наиболее целесообразных информационно-ресурсных базисов

является общий информационный фон, в первую очередь формируемый новостным потоком. В рамках единого новостного потока необходимо идентифицировать отдельный поток новостей о банке, из которых необходимо идентифицировать новости, содержащие токены, обратно ассоциирующиеся с надежности взаимодействия. В первую очередь можно выделить такие токены, как — «сбой», «сбои», «неполадки», «проблема», «мошенничество», «суд», «преступление», «утечка», «скандал» и многие другие. Данный список может масштабироваться и трансформироваться в зависимости от спецификации анализируемого сегмента ВВП-сегмента. Данный параметр может быть назван коэффициентом надежности, и рассчитан следующим образом:

$$KR_i = \frac{1}{F_{n.g.i} + 1} \quad (1)$$

Где:

1. KR_i — коэффициент надежности банка i за исследуемый период;
2. $F_{n.g.i}$ — число новостных единиц о банке i , содержащих как минимум один из выделенных ранее токенов, за исследуемый период.

Необходимо отметить, что идентификация токенов может производиться как исключительно на основе заголовков новостей, так и на основе аннотаций или полного текста. Процесс калькуляции данного коэффициента может быть автоматизирован средствами языка программирования Python 3. Данный выбор обусловлен в первую очередь наличием значительного количества библиотек для парсинга и обработки естественной цифровой информации, что в значительной мере снижает трудоемкость разработки. Алгоритм калькуляции может быть разделен на следующие последовательные этапы:

Этап 1 — Парсинг массива новостей относительно исследуемого банка. В рамках данного этапа производится автоматизированный сбор новостной информации относительно исследуемых банков. Так как результирующий коэффициент имеет относительный характер, для целей формирования конечных управленческих решений его необходимо сравнивать либо со среднерыночным значением, либо с показателями банков-конкурентов. Именно данный факт определяет необходимость сбора информационного массива сразу по нескольким банкам. Для целей автоматизации поиска и агрегирования новостной информации предлагается использо-

вать библиотеку `pygooglenews` [3]. Данный выбор обусловлен следующей спецификой:

1. Информационным источником для инструментов данной библиотеки выступают алгоритм Google News, что в значительной мере повышает универсальность получаемых результатов.

2. Данная библиотека позволяет конкретизировать как язык поиска, так и регион (страну), что позволяет ограничить новостное поле исключительно Россией и русским языком.

3. В основе данной библиотеки лежат решения на основе `selenium`, что значительно повышает стабильность поиска, и снижает вероятность блокировки со стороны Google.

Подробный алгоритм парсинга массива новостей относительно исследуемого банка представлен на рисунке 1.

Как можно видеть, данный алгоритм предполагает определение количества анализируемых банков, при последующей конкретизации из наименований. Установленные банки формирует список, для которого последовательно производится поиск новостной информации в пределах одного календарного года. Из полученной информации извлекаются исключительно заголовки. Данный выбор во много обусловлен фантомной спецификой информационного фона. Результатом данного этапа является сложный массив, представляющий собой список, состоящий из списков заголовков новостей о конкретных установленных банках за исследуемый период (календарный год).

Этап 2 — Токенизация массива новостей относительно исследуемого банка. В рамках данного этапа производится трансформация полученных ранее новостных заголовков в списки токенов. Данный этап предполагает следующие подэтапы:

1. Исключение из полученных ранее новостных заголовков знаков пунктуации.

2. Трансформация заглавных форм.

3. Первичная токенизация, предполагающая трансформацию единых новостных заголовков в списки последовательных лексем (токенов).

4. Лемматизация полученных токенов, для целей формирования списков, содержащих исключительно лексемы в базовой словарной форме.

5. Исключение из списков лемматизированных токенов лексем с низкой содержатель-

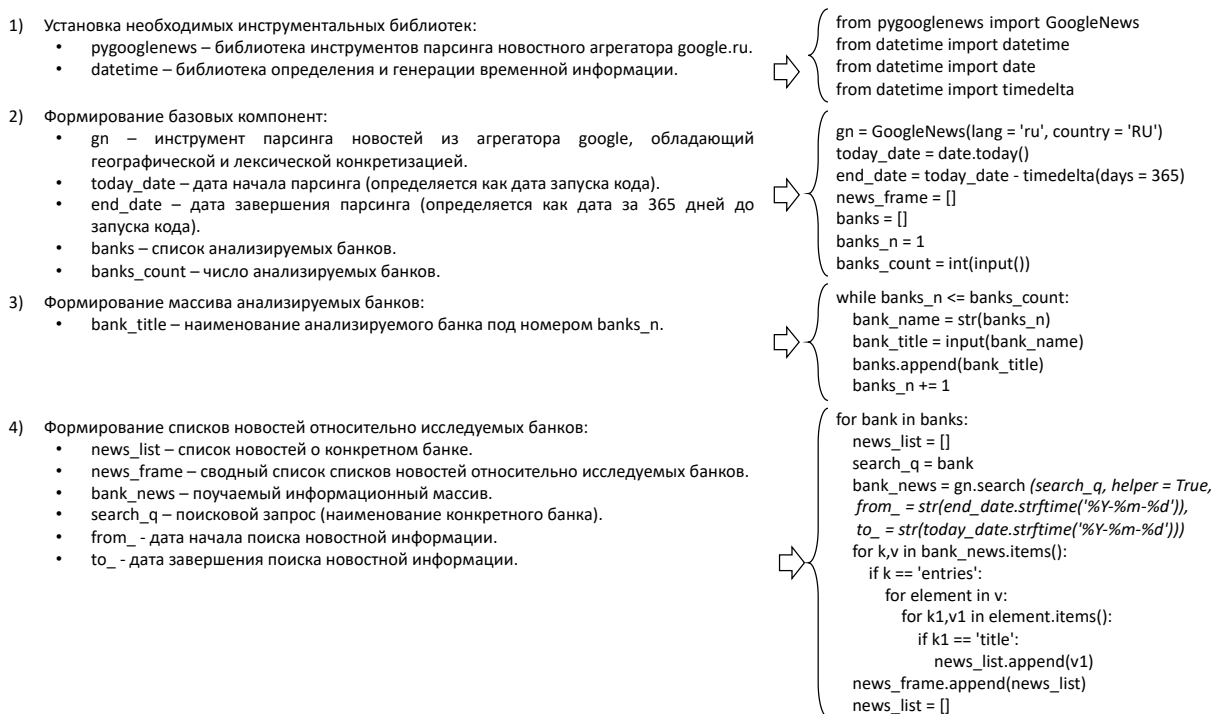


Рисунок 1. Автоматизированный алгоритм парсинга массива новостей относительно исследуемого банка

ной значимостью (предлогов, союзов, местоимений и тд).

Подробный алгоритм токенизации массива новостей относительно исследуемого банка представлен на рисунке 2.

По результатам реализации данного этапа формируется сложный массив, представляющий собой список, состоящий из списков токенов, описывающих новостной фон конкретного анализируемого банка за исследуемый период (календарный год).

Этап 3 – Идентификация новостей, содержащих как минимум один из токенов, обратно ассоциирующиеся с надежности взаимодействия с банком. В рамках данного этапа производится сравнение сформированных на предыдущем этапе списков токенов с определенным списком токенов, обратно ассоциирующихся с надежностью взаимодействия с банком. Данный список также был приведен ранее, однако, необходимо отметить, что он может корректироваться в зависимости от условий конкретизации сегмента анализируемого VIP-сегмента потребителей. Подробный алгоритм идентификации новостей, содержащих как минимум один из токенов, обратно ассоциирующиеся с надежности взаимодействия с банком представлен на рисунке 3.

Результатам реализации данного этапа является список, содержащий введенное ранее число новостных единиц, содержащих как минимум один из соответствующих токенов, для каждого из исследуемых банков за исследуемый период (календарный год). Однако, данной показатель необходимо инвертировать, а также нормализовать относительно сформированного массива. Операция нормализации может трансформироваться, однако в рамках данного примера предлагается рассмотреть наиболее универсальное решение, не требующее введения нормативных значений исследуемых переменных.

Этап 4 – Расчет и нормализация коэффициента надежности. Данный этап является завершающим в рамках алгоритма калькуляции коэффициента надежности взаимодействия с банком, и он предполагает последовательное инвертирование и нормализацию полученных на предыдущем этапе значений. Данный алгоритм представлен на рисунке 4.

По результатам реализации данного этапа формируется список нормализованных коэффициентов надежности для каждого из исследуемых банков. Данный коэффициент позволяет сравнивать исследуемые банки и принимать на основе данного сравнения управленческие решения. Расчет данного коэффициента самим

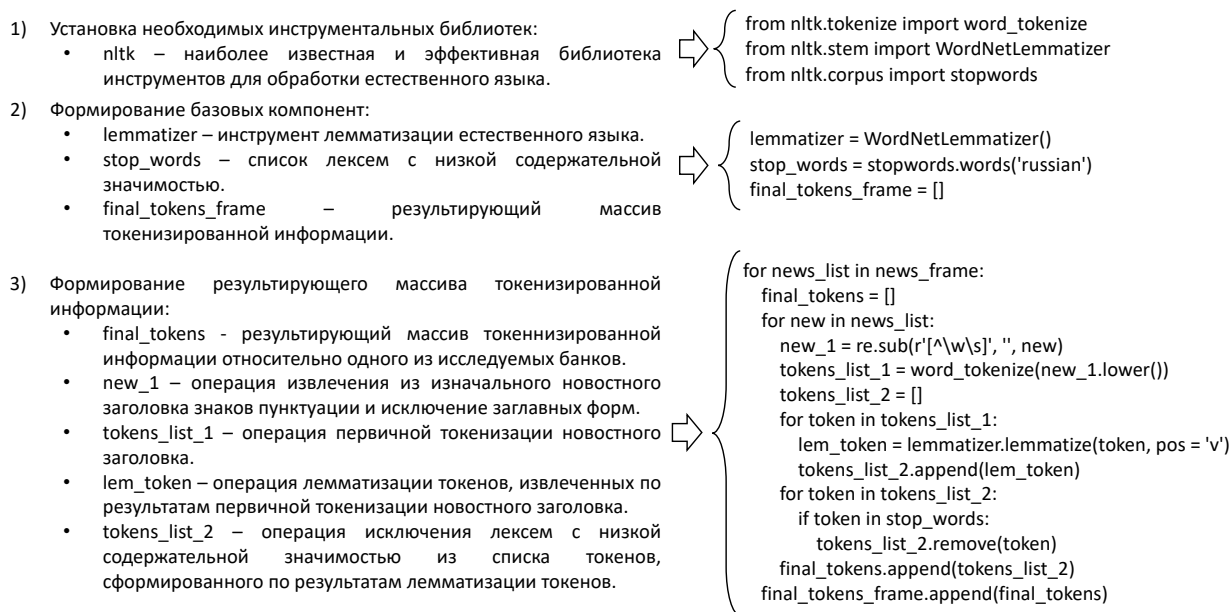


Рисунок 2. Автоматизированный алгоритм токенизации массива новостей относительно исследуемого банка

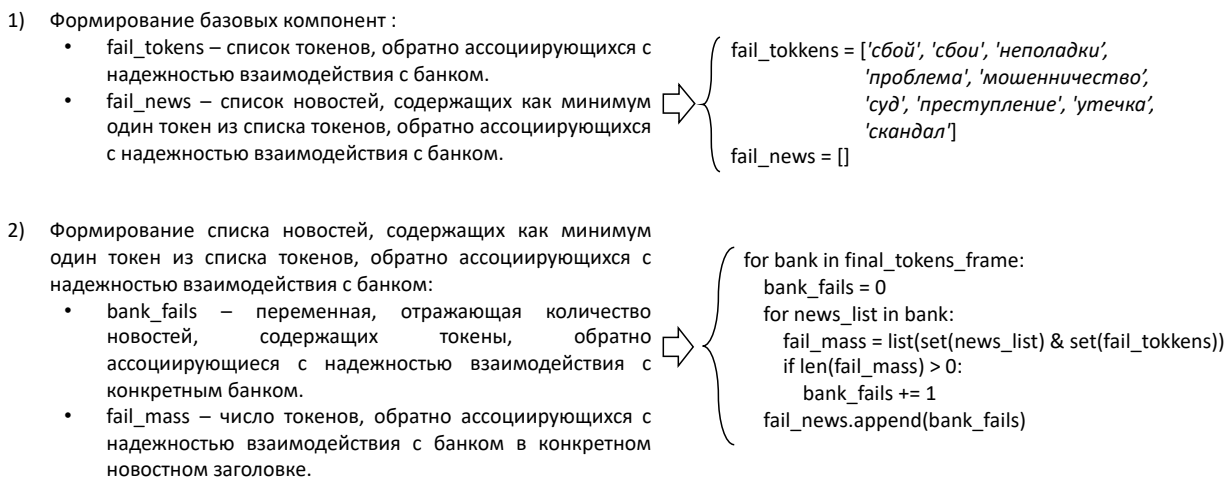


Рисунок 3. Автоматизированный алгоритм идентификации новостей, содержащих как минимум один из токенов, обратно ассоциирующиеся с надежности взаимодействия с банком

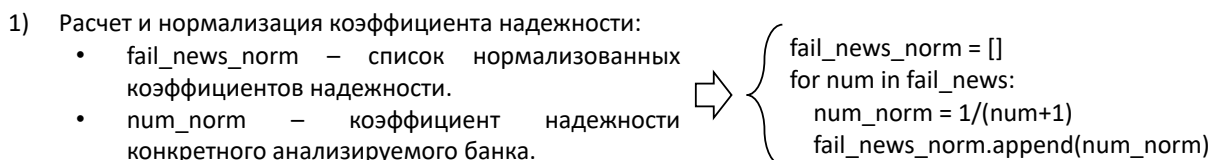


Рисунок 4. Автоматизированный алгоритм расчета и нормализация коэффициента надежности взаимодействия с банком

банком может быть частью конкурентного анализа.

Далее необходимо рассмотреть совместимость восприятия банка. Данное свойство отражает как содержательное так и эмоциональное восприятие банка. Однако, подавляющий массив содержательных компонент является единым для большинства банков. Данное утверждение основано на единообразии оказываемых услуг. Следовательно, наиболее целесообразно интерпретировать совместимость восприятия банка исключительно на эмоциональном уровне. Данный уровень в преломлении методологии обработки естественной цифровой информации может быть проанализирован посредством инструментов оценки тональности. Тональность или тон текста — это отражение эмоционального отношения автора к каким-либо событиям или объектам (эмоциональная составляющая текста) [4]. В науке выделяет множество тональных характеристик текста, однако применительно к представителям ВИП-сегмента потребителей можно выделить такие характеристики как уровень нейтральности, уровень позитивности, уровень негативности, уровень лексической упрощенности и уровень средовой единообразности. Соотношение данных характеристик уникально для каждого из возможных сегментов потребителей банковских услуг. Следовательно, коэффициент, отражающий значение данного свойства для исследуемого банка может быть назван коэффициентом тональной совместимости. Основываясь на результатах проведенного ранее исследования, представленного в статье [1], можно утверждать, что аватар представителя ВИП-сегмента потребителей банковских услуг склонен отдавать предпочтение банкам, характеризующимся максимальным значением следующего соотношения тональных компонент:

$$KT_i = \frac{\overline{T_{neutral_i}} + \overline{T_{skip_i}}}{\overline{T_{positive_i}} + \overline{T_{negative_i}} + \overline{T_{speech_i}}} \quad (2)$$

Где:

1. KT_i — коэффициент тональной совместимости банка i за исследуемый период;
2. $\overline{T_{neutral_i}}$ — средний уровень нейтральности анализируемого контента, представляющего банк i , за исследуемый период;
3. $\overline{T_{skip_i}}$ — средний уровень средовой единообразности анализируемого контента, представляющего банк i , за исследуемый период;

4. $\overline{T_{positive_i}}$ — средний уровень позитивности анализируемого контента, представляющего банк i , за исследуемый период;

5. $\overline{T_{negative_i}}$ — средний уровень негативности анализируемого контента, представляющего банк i , за исследуемый период;

6. $\overline{T_{speech_i}}$ — средний уровень лексической упрощенности анализируемого контента, представляющего банк i , за исследуемый период.

Как выделялось ранее, анализируемый контент, представляющий банк, крайне дифференцирован. В рамках данного исследования выдвигается гипотеза, согласно которой наиболее значимым и релевантным для тонального анализа выступает контент, генерируемый непосредственно клиентами банка. Данная гипотеза основывается на предположении о том, что банк в первую очередь стремится к удовлетворению потребностей основных потребителей, что предполагает адаптацию коммуникативных паттернов. В том случае, если основная аудитория потребителей услуг конкретного банка характеризуется генерацией естественной цифровой информации, в меньшей степени соотносящейся с выделенным соотношением тональных характеристик, следовательно с высокой долей вероятности коммуникативные паттерны данного банка будут в меньшей степени удовлетворять представителей ВИП-сегмента потребителей. Следовательно, в качестве аналитического базиса может быть использован массив естественной цифровой информации, генерируемой потребителями услуг конкретного банка, в социальных сетях (в частности в комментариях в официальных сообществах банков). Таким образом, алгоритм калькуляции коэффициента тональной совместимости может быть представлен следующим последовательными этапами:

Этап 1 — Парсинг идентификаторов записей из официального сообщества банка в социальной сети. В качестве анализируемой социальной сети в рамках данного исследования предлагается использовать «ВКонтакте». Данный выбор обусловлен следующей спецификой:

1. Данная социальная сеть является исключительно отечественной, что определяет национальную принадлежность пользователей.
2. Данная социальная сеть является одной из крупнейших в интернете, в связи с чем подавляющее число банков имеет официальные сообщества в данной социальной сети.
3. Данная социальная сеть обладает бес-

платным API интерфейсом, позволяющим автоматизировать процесс парсинга информации, содержащейся в ней.

Необходимо отметить, что процесс парсинга информации из официального сообщества банка в социальной сети «ВКонтакте» является в достаточной мере универсальным, в связи с чем часть алгоритмов дублируется. Подробный алгоритм парсинга идентификаторов записей из официального сообщества банка в социальной сети представлен на рисунке 5.

Результатом реализации данного этапа является сложный массив, состоящий из ключей в форме наименований банков, и значений, сохраненных в данных ключах, в форме списков идентификаторов записей из официальных сообществ анализируемых банков в социальной сети «ВКонтакте». Данный массив является промежуточным результатом, и необходим исключительно для формирования массивов комментариев пользователей к данным постам.

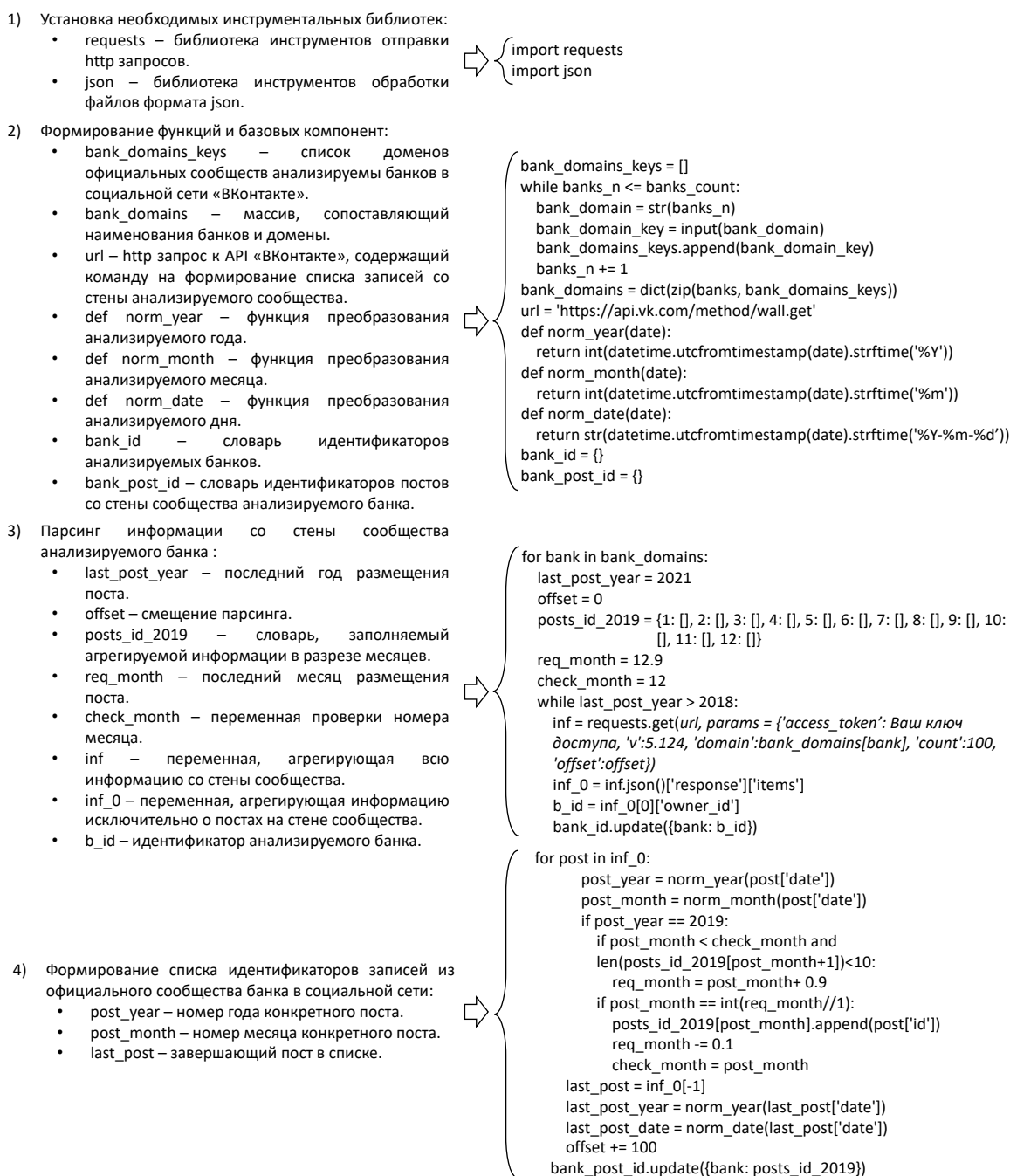


Рисунок 5. Автоматизированный алгоритм парсинга идентификаторов записей из официального сообщества банка в социальной сети

Этап 2 — Парсинг комментариев пользователей к записям из официального сообщества банка в социальной сети. На основе сформированного на предыдущем этапе массива идентификаторов постов, в рамках данного этапа производится последовательное агрегирование комментариев пользователей к данным постам. В случае отсутствия комментариев массив заполняется нулевыми значениями. Подробный алгоритм парсинга комментариев пользователей к записям из официального сообщества банка в социальной сети представлен на рисунке 6.

Результатом реализации данного этапа является структурированный массив комментарием, определенный для каждого анализируемого банка. Далее, производится оценка тональных компонент каждого из агрегированных комментарием.

Этап 3 — Оценка тональности комментариев пользователей к записям из официального сообщества банка в социальной сети. Для целей реализации данного этапа используется специализированная библиотека инструментов оценки тональности текстов на русском языке «Dostoevsky». Выбор данной библиотеки обусловлен следующей спецификой:

1. Библиотека «Dostoevsky» является единственной эффективной библиотекой Python, позволяющей анализировать все из выделенных ранее тональных компонент текста.

2. Библиотека «Dostoevsky» является единственной общедоступной библиотекой Python,

позволяющей анализировать тональность текстов на русском языке.

3. Библиотека «Dostoevsky» содержит собственные инструменты токенизации, что значительно упрощает процесс предварительной подготовки текстов.

В рамках данного этапа в первую очередь производится оценка тональных компонент каждого комментария в отдельности, после чего производится усреднение полученных значений для каждого из анализируемых банков. Подробный алгоритм оценки тональности комментариев пользователей к записям из официального сообщества банка в социальной сети представлен на рисунке 7.

По результатам реализации данного этапа формируется массив данных, содержащий усредненные значения необходимых тональных компонент комментариев к постам в официальных сообществах анализируемых банков. В рамках завершающего этапа необходимо произвести калькуляцию коэффициентов тональной совместимости для каждого из анализируемых банков.

Этап 4 — Расчет коэффициента тональной совместимости. Коэффициент тональной совместимости калькулируется для каждого из анализируемых банков. Так как получаемые значения тональных компонент измеряются от 0 до 1, нормализация конечного коэффициента не является необходимой. Алгоритм расчёта коэффициента тональной совместимости представлен на рисунке 8.

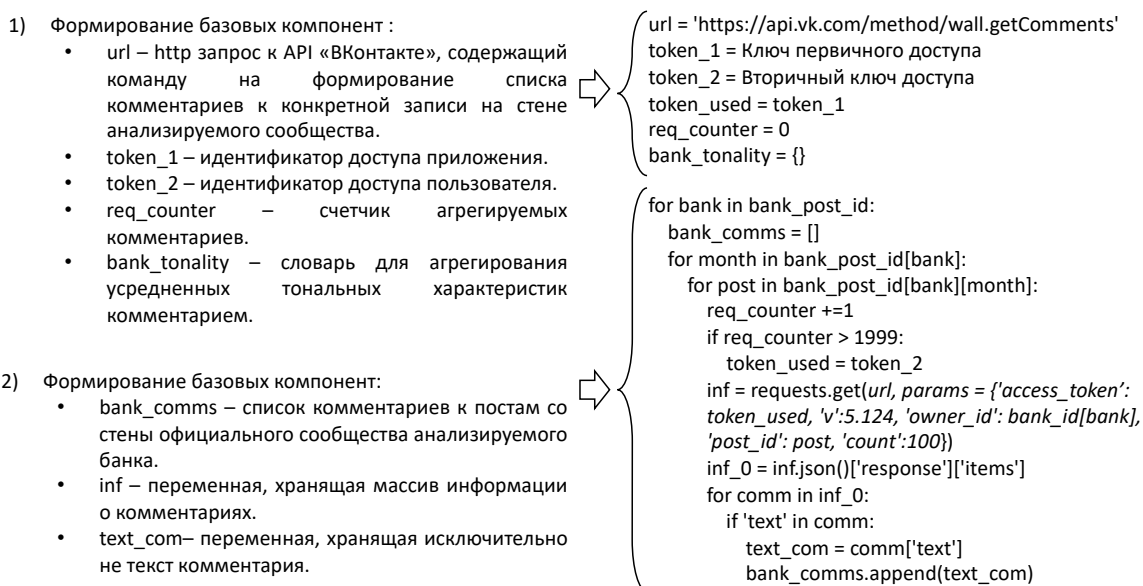


Рисунок 6. Автоматизированный алгоритм парсинга комментариев пользователей к записям из официального сообщества банка в социальной сети

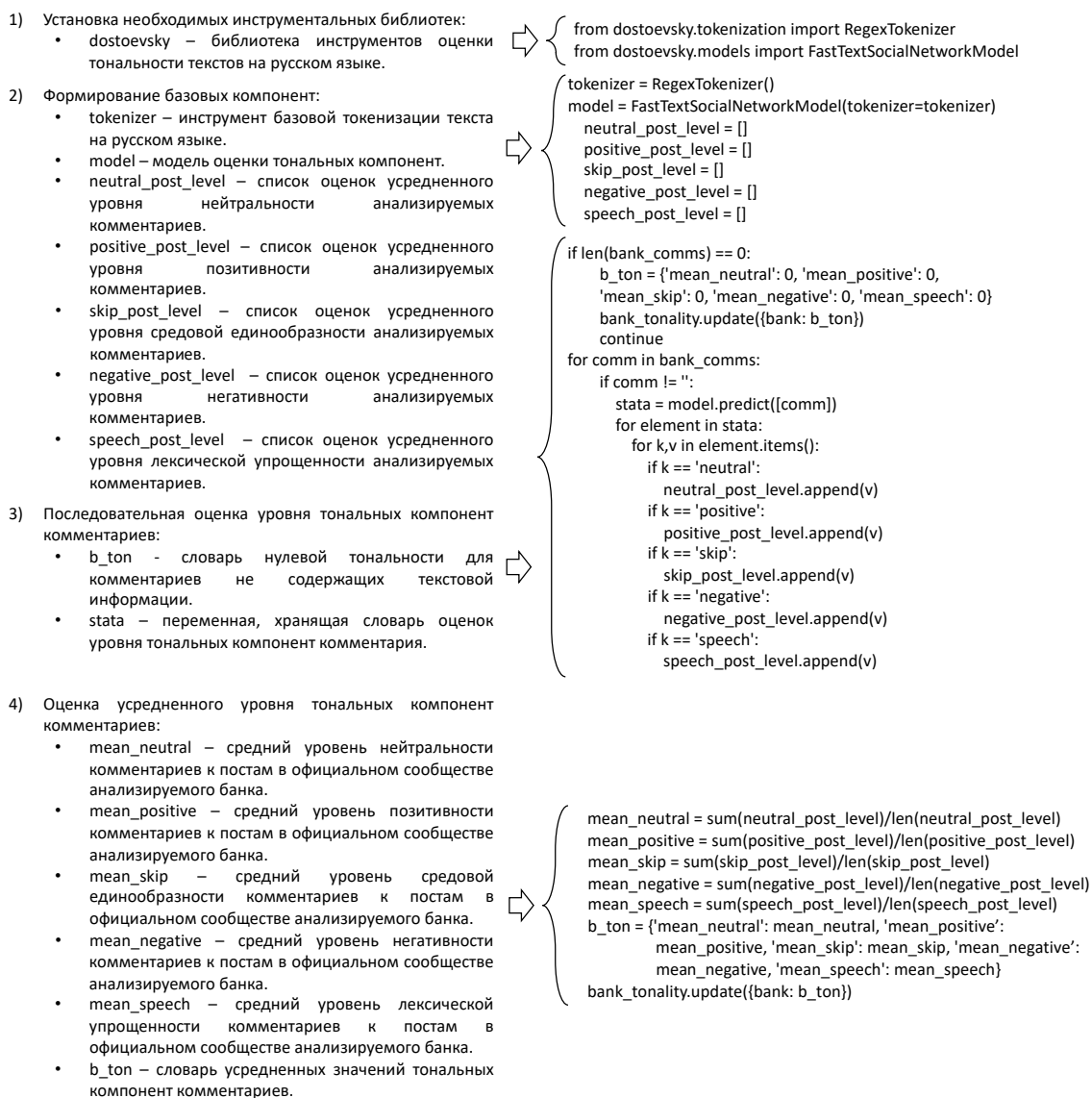


Рисунок 7. Автоматизированный алгоритм оценки тональности комментариев пользователей к записям из официального сообщества банка в социальной сети

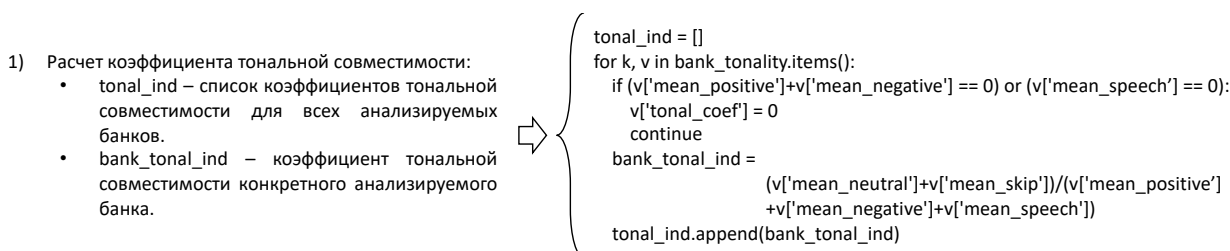


Рисунок 8. Автоматизированный алгоритм расчета коэффициента тональной совместимости для анализируемых банков

По результатам реализации данного этапа формируется список коэффициентов тональной совместимости для каждого из исследуемых банков. На основе сравнительного анализа значений данных коэффициентов также могут уточняться управленческие решения и производится конкурентный анализ. В сочетании с коэффициентом надежности, данный коэффициент формирует условно-традиционный ценностно-сравнительный аналитический результат для представителей ВИП-сегмента потребителей банковских услуг. Однако, по результатам динамичного процесса цифровизации все большую значимость приобретает сравнительная оценка следующего параметра, отражающего инновационную активность банка.

В завершении необходимо рассмотреть квантификатор инновационной активности банка, в контексте естественного восприятия. На данный момент в науке существует множество показателей, отражающих инновационную активность предприятия. Однако, в данном случае фактическая инновационная активность значима в меньшей степени. В данном исследовании предлагается квантифицировать восприятие инновационной активности банка, посредством анализа общего информационного фона. Данный показатель может быть назван — коэффициент инновационной ассоциативности. Сущность данного коэффициента определяется долей информационного фона банка, содержащей информацию относительно его инновационных достижений. Подавляющая часть инновационных достижений банка сконцентрировано в области «Финтех», так как данная область концентрирует в себе все инновационные достижения в области финансовых технологий. Следовательно, тематические информационные сообщества являются концентраторами токенов, характеризующих данную область. Таким образом, для целей калькуляции коэффициента инновационной ассоциативности банк можно использовать принцип, который был заложен в процедуру квантификации коэффициента надежности, описанного ранее, а именно — доля токенов, характеризующих область «Финтех», в общем массиве токенов, характеризующих информационный фон банка, отражает инновационную активность анализируемого банка.

$$KIn_i = \frac{F_{in_i}}{F_i} \quad (3)$$

Где:

1. KIn_i — коэффициент инновационной ассоциативности банка i за исследуемый период;
2. F_{in_i} — доля токенов, характеризующих область «Финтех»;
3. F_i — общая совокупность токенов, описывающих информационный фон банка i за исследуемый период.

Необходимо учитывать, что область «Финтех» является крайне динамичной, и массив токенов ее описывающий также непостоянен. Следовательно, целесообразно использовать обновляющийся массив, для описания данной области. Для целей формирования совокупности токенов, описывающих информационный фон банка, могут быть использованы посты в официальном сообществе банка в социальной сети. Таким образом, для целей автоматизированной калькуляции коэффициента инновационной ассоциативности банка может быть использован следующий алгоритм:

Этап 1 — Парсинг информации, актуально характеризующей область «Финтех». Данный этап технологически близок к этапу 2 алгоритма калькуляции коэффициента тональной совместимости. В качестве концентратора тематической информации в области «Финтех» предлагается использовать сообщество в социальной сети «ВКонтакте» — «fintechlabrussia». Данный выбор в первую очередь обусловлен соответствием необходимой тематики, а также периодическим обновлением материалов. Источником необходимой информации выступают посты на стене анализируемого сообщества. Автоматизированный алгоритм агрегирования текстов постов со стена анализируемого сообщества представлен на рисунке 9.

По результатам данного этапа формируется список, содержащий текст постов, посвященных области «Финтех». Далее, данные тексте требуют токенизации.

Этап 2 — Токенизация информации, актуально характеризующей область «Финтех». Алгоритмы токенизации были подробно рассмотрены ранее. Однако в данном случае необходимо отметить необходимость квантификации токенизированной информации в форме частоты упоминания. Данный этап необходим для последующего определения удельного веса каждого токена. Автоматизированный алгоритм токенизации текстов постов со стена анализируемого сообщества представлен на рисунке 10.

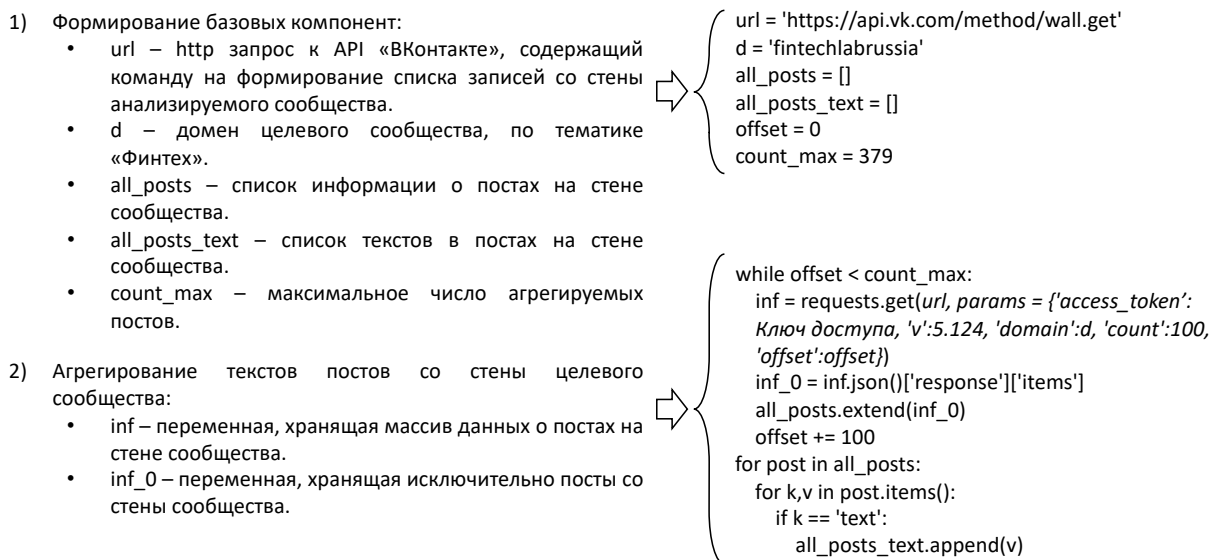


Рисунок 9. Автоматизированный алгоритм агрегирования текстов постов со стена анализируемого сообщества

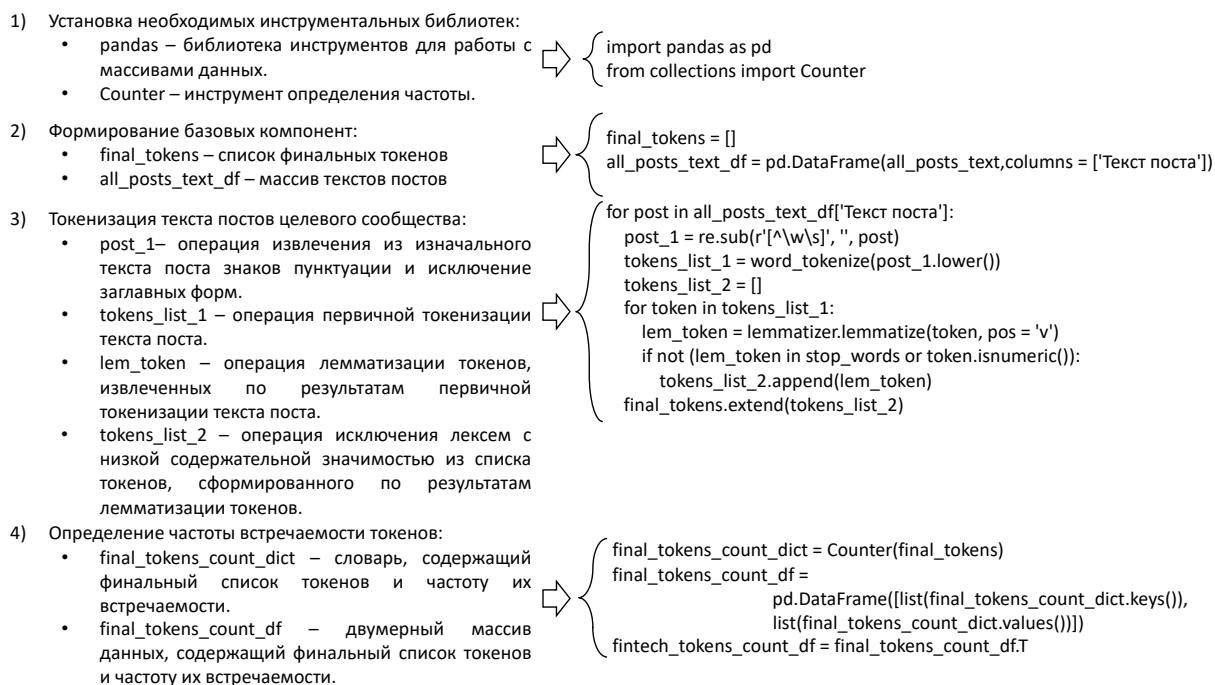


Рисунок 10. Автоматизированный алгоритм токенизации текстов постов со стена анализируемого сообщества

По результатам данного этапа формируется массив данных, содержащий как сами токены, характеризующие область «Финтех», так и их частоту, необходимую для определения удельного веса каждого токена. Для целей реализации операция сравнения далее необходимо продублировать описанные этапы для официального сообщества анализируемого банка

Этап 3 – Парсинг информации, актуально описывающей информационный фон анализируемого банка. По результатам реализации данного этапа формируется аналитическая матрица, содержащая посты со стены официального сообщества анализируемого банка в социальной сети. Автоматизированный алгоритм агрегирования информации, актуально описывающий информационный фон анализируемого банка представлен на рисунке 11.

Получаемая по результатам данного этапа аналитическая матрица далее требует токенизации.

Этап 4 – Токенизация информации, актуально описывающей информационный фон анализируемого банка. Данный этап технологически идентичен этапу 2 данного алгоритма. По результатам токенизации формируется массив данных, содержащий как сами токены, описывающие информационный фон анализируемого банка, так и их частоту, необходимую для определения удельного веса каждого токена. На завершающем этапе данного алгоритма произ-

водится расчет коэффициента инновационной ассоциативности анализируемого банка.

Этап 5 – Расчет коэффициента инновационной ассоциативности банка. Автоматизированный алгоритм расчета коэффициента инновационной ассоциативности анализируемого банка представлен на рисунке 12.

По результатам калькулирования формируется массив коэффициентов инновационной ассоциативности, соответствующих анализируемым банкам. Для целей комплексного сравнительно анализа профилей привлекательности банка для представителей ВИП-сегмента потребителей банковских услуг, наиболее целесообразно использовать лепестковую диаграмму. Предложенный алгоритм был применён по отношению к 3 банкам: «Промсвязьбанк», «Банк Открытие» и «Сбербанк». На рисунке 13 представлены цифровые профили привлекательности банков для представителей ВИП-сегмента потребителей.

Как можно видеть, доминантой для всех анализируемых банков является коэффициент инновационной активности, в то время как иные коэффициенты значительно ниже. Однако, на сравнительно более привлекательным для представителей ВИП-сегмента потребителей является «Промсвязьбанк», что в первую очередь обеспечивается коэффициентом тональной совместимости.

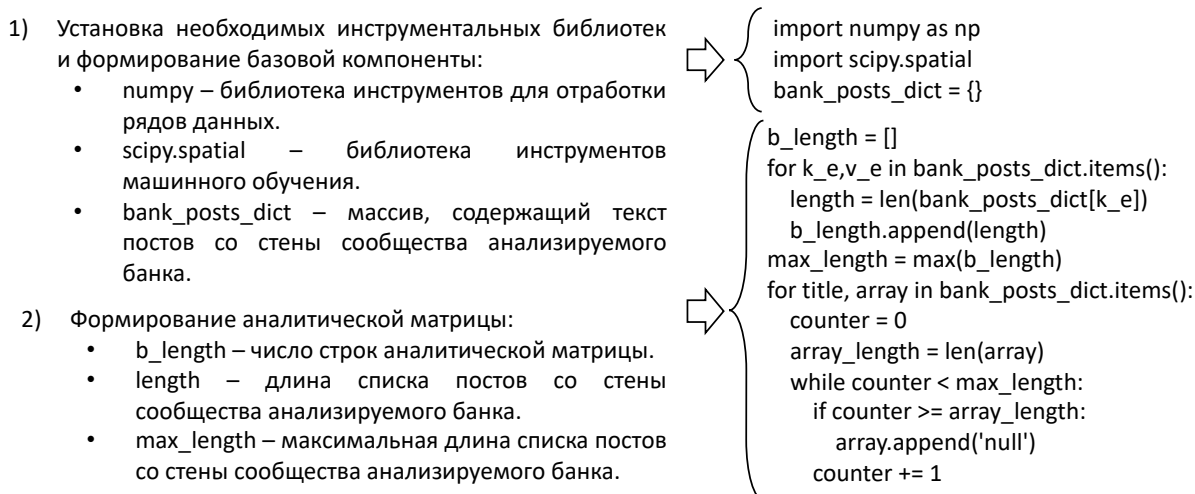


Рисунок 11. Автоматизированный алгоритм описывающий информационный фон анализируемого банка

1) Формирование базовых компонент:

- `b_tokens_dict` – словарь, объединяющий токены, присутствующие в обоих аналитических массивах.
- `tokens_count_dict` – словарь, содержащий перекрестную частоту токенов, присутствующие в обоих аналитических массивах.
- `ft_frame` – двумерный массив, объединяющий токены, присутствующие в обоих аналитических массивах и их частоту.

```
b_tokens_dict = {}
tokens_count_dict = Counter(b_tokens)
b_tokens_dict[bank] = tokens_count_dict
ft_frame = fintech_tokens_count_df.T
tokens = ft_frame.iloc[0].to_list()
freq = ft_frame.iloc[1].to_list()
```

2) Определение коэффициентов инновационной ассоциативности:

- `word_list` – список токенов, присутствующие в обоих аналитических массивах.
- `rasst_list` – коэффициент инновационной ассоциативности анализируемого банка.
- `bank_rasst_list` – список коэффициентов инновационной ассоциативности анализируемых банков

```
word_list = []
for token in tokens:
    if token not in word_list:
        word_list.append(token)
for bank in b_tokens_dict:
    for token in b_tokens_dict[bank]:
        if token not in word_list:
            word_list.append(token)
tok_dict = dict(zip(range(len(word_list)), word_list))
a = np.zeros((len(b_tokens_dict)+1, len(tok_dict)))
i = 0
j = 0
for token in freq:
    a[i, j] = token
    j += 1
i += 1
for bank in b_tokens_dict:
    for token in b_tokens_dict[bank]:
        j = word_list.index(token)
        a[i, j] = b_tokens_dict[bank][token]
i += 1
u = a[0]
rasst_list = []
line_number = 0
for v in a:
    rasst = scipy.spatial.distance.cosine(u, v)
    line_number += 1
    rasst_list.append(rasst)
bank_rasst_list = rasst_list[1:]
```

Рисунок 12. Автоматизированный алгоритм расчета коэффициента инновационной ассоциативности анализируемого банка

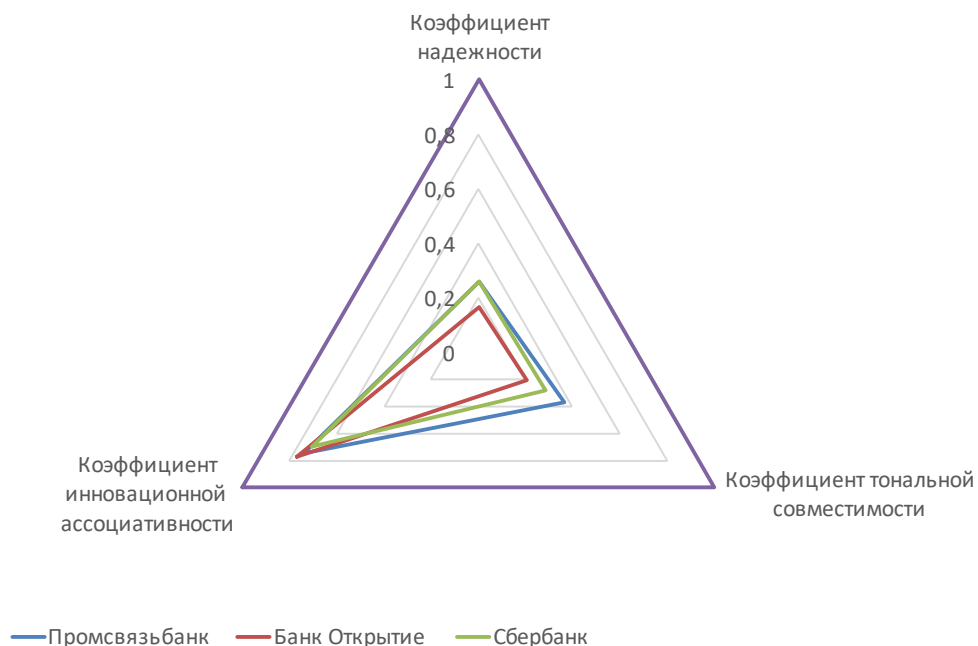


Рисунок 13. Цифровые профили привлекательности анализируемых банков для представителей ВИП-сегмента потребителей

Библиографический список

1. *Ершова А.В., Шубаева В.Г.* Применение теории аватаров в банковском маркетинге ВВП-сегмента // *Экономические науки.* 2020. № 193.
2. *Словарь русского языка: В 4-х т. / РАН, Ин-т лингвистич. исследований; Под ред. А. П. Евгеньевой. — 4-е изд., стер. — М.: Рус. яз.; Полиграфресурсы, 1999.*
3. Pygooglenews // Github URL: <https://github.com/kotartemiy/pygooglenews> (дата обращения: 25.01.2021).
4. Что такое тональность текста и каким образом ее создать // Блог Думай и пиши URL: <https://writingforonline/post/tonalnost-teksta/> (дата обращения: 25.01.2021).
5. *Конников Е.А., Терентьева Д. А., Конникова О. А.* Анализ уровня устойчивого потребления в контексте цифровой информационной среды // *Экономические науки.* 2020. № 192. С. 114–125.
6. *Родионов Д.Г., Конников Е. А., Алферьев Д. А.* Информационный капитал предприятия как целевой показатель развития в рамках цифровых экономических систем // *Экономические науки.* 2020. № 190. С. 131–137.