

СТАТЬЯ

УДК 004

**СЕРВИС ПОПУЛЯРНОСТИ АРТИСТОВ
НА ОСНОВЕ АНАЛИЗА СОЦИАЛЬНЫХ СЕТЕЙ****Хуснияров И.Ф.***ООО «Тур», Уфа, e-mail: Tour@fangid.com*

Сегодня социальные сети играют огромную роль в жизни каждого человека, и с каждым днем пользователей социальных сетей становится все больше. При этом социальные сети выступают не только как площадка для общения и поиска друзей, но и средство для прослушивания треков, просмотра видеозаписей концертов, знакомства с артистами, выражения мнения относительно их творчества. Поэтому сложной задачей становится определение популярности артистов, а имеющиеся модели оценки популярности, основанные на биографических данных, статусах, наградах и радиочартах, теряют свою актуальность. В статье приведены результаты исследования популярности музыкальных артистов на основе математического моделирования и анализа социальных сетей. В основу анализа заложены более 20 источников данных, 86 метрик. Результаты моделирования апробированы на 6000 артистов России. Полученные результаты агрегируются и представляются на разработанной платформе музыкальной аналитики, которая позволяет увидеть и отслеживать динамику популярности, рейтинговать исполнителей, а также увидеть популярность артистов в региональном разрезе. Данная платформа может являться основой для проработки рекламных кампаний, определения музыкальных трендов, формирования музыкальных чартов на стриминговых сервисах и организации концертной деятельности.

Ключевые слова: популярность артиста, рейтинг, сервис, социальные сети, анализ данных, музыкальная аналитика

**ARTIST POPULARITY SERVICE BASED
ON SOCIAL NETWORKS ANALYSIS****Khusniyarov I.F.***Tour LLC, Ufa, e-mail: Tour@fangid.com*

Today, social networks play a huge role in the life of every person, and every day there are more and more users of social networks. At the same time, social networks act not only as a platform for communication and finding friends, but also as a means for listening to tracks, watching videos of concerts, meeting artists, expressing opinions about his work. Therefore, determining the popularity of artists becomes a difficult task, and the existing models for assessing popularity based on biographical data, status, awards and radio charts are losing their relevance. The article presents the results of a study of the popularity of musical artists based on mathematical modeling and analysis of social networks. The analysis is based on more than 20 data sources, 86 metrics. The simulation results have been tested on 6,000 artists in Russia. The results obtained are aggregated and presented on the developed music analytics platform, which allows you to track the dynamics of popularity, rank performers, and also see the popularity of artists in the regional context. This platform can be the basis for developing advertising campaigns, identifying music trends, forming music charts on streaming services, and organizing concert activities.

Keywords: artist popularity, rating, service, social networks, data analysis, music analytics

Стремительное развитие социальных сетей за последние несколько лет связано в первую очередь с технологическими прорывами и пандемией. Подобное развитие технологий детерминировало медиасреду и медиапотребление [1]. Согласно данным статистики [2], социальные сети используют более 4,5 млрд чел., а темпы прироста пользователей за 2021 г. составили почти 10%.

Самыми популярными социальными сетями являются Youtube (более 90 млн пользователей в России), ВКонтакте (более 70 млн пользователей в России), TikTok (более 35 млн пользователей в России) [3]. При этом для всех социальных сетей стоит отметить высокую вовлеченность пользователей: соцсети не только площадка для общения и поиска друзей, но и средство для прослушивания треков, просмотра ви-

деозаписей концертов, знакомства с артистами, выражения мнения относительно их творчества. Одним из результатов развития данного направления стало активное вовлечение артистов «старой формации» (получивших признание, популярность среди населения до появления социальных сетей и развития интернета) в социальные сети, а также появление артистов «новой формации» («выходцев» из социальных сетей).

В связи с этим для моделирования оценки популярности артистов в качестве основы были взяты социальные сети, стриминговые сервисы, радио, чарты и площадки для просмотра клипов. Отдельным направлением стало исследование популярности артистов на основе динамики запросов в поисковых сервисах.

Данные по каждому из параметров собираются с помощью парсеров и агреги-

руются в показатели рейтинга по каждому из параметров. На основе интегрального показателя формируется итоговый рейтинг популярности артиста.

Цель формирования рейтинга артистов на основе их популярности в социальных сетях заключается в формировании единого подхода к оценке популярности с учетом развития информационно-коммуникационных технологий. Непосредственными задачами при этом являются:

- 1) анализ существующих методов и показателей для рейтингования;
- 2) анализ методов сбора и учета данных;
- 3) разработка математической модели исследования;
- 4) разработка цифровой платформы музыкальной аналитики.

Основным принципом при создании рейтинга и методики оценки стало использование электронных форм сбора информации. Значимыми особенностями для разработки математической модели стали учет специфики каждой социальной сети и фиксирование текущей популярности артиста. Расчет рейтинга производится путем прямого сравнения показателей. Информационной основой – данные по артистам на основе анализа их страниц в социальных сетях и статистике запросов слушателей стриминговых сервисов, радиостанций, пользователей сети Интернет.

Алгоритм проведения исследования

Для оценки и рейтингования артистов по их популярности в социальных сетях был разработан алгоритм оценки, состоящий из нескольких этапов.

1. На первом этапе выбирается периодичность данных (день/неделя/месяц/квартал/год). В рамках сервиса музыкальной аналитики данный этап представлен выбором дат.

2. На втором этапе проводится расчет разницы между метриками выбранной социальной сети между значениями на начало и конец периода.

3. На третьем этапе проводится математическое моделирование. Для каждой социальной сети была разработана математическая формула определения популярности.

4. На четвертом этапе формируется рейтинг артистов по популярности на основе прямого ранжирования на основе показателя, полученного на третьем этапе.

5. На пятом этапе формируется итоговый рейтинг артиста на основе интегрального показателя рейтингов в различных социальных сетях с учетом популярности и распространенности социальной сети на территории Российской Федерации.

В рамках работы сервиса музыкальной аналитики реализовано визуальное представление рейтингов артистов как для отдельно выбранной социальной сети, так и интегральный рейтинг на основе всех параметров.

Статистическая информация, необходимая для исследования популярности артистов

Для оценки популярности артистов в социальных сетях используются различные метрики, полученные на основе анализа страниц артистов в социальных сетях: количество размещенных постов, видео, треков, а также реакций пользователей и подписчиков на информационный материал, выраженный в комментариях, лайках, дизлайках, просмотрах, репостах и прослушиваниях. Для рейтингования по динамике запросов используются количественные показатели. Для рейтингования на основе данных музыкальных чартов используются дополнительные параметры: наивысшая позиция артиста в чарте, дата получения этой позиции, количество дней. Ключевые показатели и метрики для оценки популярности в социальных сетях представлены в таблице.

Сбор необходимой статистической информации проводится с использованием программы сбора и систематизации информации в несколько этапов.

1. На первом этапе проводится сканирование исходного массива информации страницы в социальных сетях, базы данных стримингового сервиса, чарта, запросов.

2. На втором этапе проводится выделение значимых показателей по заданным параметрам.

3. На третьем этапе проводится конвертация полученных данных в необходимый формат и агрегация полученных результатов в разработанной базе данных.

Проверка качества исходной информации

Собираемые для проведения исследования данные должны отвечать определенным требованиям:

- 1) достоверности – соответствию данных тому, что есть на самом деле. В настоящем исследовании методика, техника и организация проведения статистического наблюдения направлены на обеспечение достоверных данных. Как известно, общим условием обеспечения достоверности является полнота охвата наблюдаемого объекта, то есть полнота и точность регистрации данных по каждой единице наблюдения [4]. Это условие выполняется на основе обновления данных в режиме реального времени.

Ключевые метрики социальных сетей

Социальная сеть/Сервис	Метрики
ВКонтакте	Количество записей артиста Количество репостов записей Количество лайков Количество комментариев Количество просмотров Количество подписчиков
Tik-Tok	Количество видео с треками артиста Количество лайков Количество комментариев Количество просмотров Количество подписчиков
YouTube	Количество видео артиста Количество дизлайков Количество лайков Количество комментариев Количество просмотров Количество подписчиков
Wordstat и GoogleAds, Википедия, Shazam	Количество запросов
Радиостанции	Количество треков артиста на радиостанции Количество воспроизведений в эфире
Чарты (Яндекс Музыка, Spotify, VK, Youtube, Apple Music)	Количество дней в чарте Средняя позиция в чарте Дата наивысшей позиции в чарте
Стриминговые сервисы	Количество подписчиков плейлистов Средняя позиция плейлиста Количество плейлистов Количество подписчиков артиста

2) возможности обобщения данных об отдельных явлениях или их сопоставимости друг с другом, то есть чтобы данные собирались в одно и то же время и по единой методике. Для выполнения данного условия все показатели должны быть приведены к стандартизированному виду, что обеспечивается работой программы сбора и систематизации информации перед занесением в базу данных.

Постановка задачи определения весовых коэффициентов социальных сетей, радиостанций и поисковых сервисов

Учитывая распространенность и зависимость некоторых параметров от используемого аппаратного обеспечения и личных предпочтений пользователей, отдельной задачей становится определение весовых коэффициентов социальных сетей, радиостанций и поисковых сервисов. Для получения значений весовых коэффициентов используется алгоритм нахождения медианного значения на основе процентного соотношения показателей.

Согласно данному алгоритму для каждого артиста суммируются значения по заданным показателям (запросы, количества треков и т.д.), затем рассчитывается долевое

соотношение по каждому поисковому сервису или радиостанции. Для полученного ряда долей по всем артистам формируется медианное значение, которое становится основой для определения весового коэффициента.

Дальнейшее исследование проводится в два этапа. На первом этапе из имеющейся совокупности показателей $\{x_1, x_2, x_3, \dots, x_n\}$ определяется $y_i = (x_1, x_2, \dots, x_n)$, где y_i – долевое значение по отобранным показателям артиста в i -й социальной сети.

На втором этапе на основе полученных данных определяется $R_i = f(y_1, y_2, \dots, y_m)$, где R_i – итоговый весовой коэффициент для каждой социальной сети.

Инструменты сервиса музыкальной аналитики и рейтингования артистов на основе их популярности в социальных сетях

Разработанный сервис собирает из открытых источников данных статистику об артистах, концертах, альбомах, треках, клипах по заданным метрикам. Полученные данные позволяют оценить популярность того или иного артиста, следить за трендами, сравнивать каналы продвижения артистов. Для отдельных пользователей сервиса возможно получение собранных данных

на основе закрытого API для последующего моделирования.

Основным инструментом сбора данных служат парсеры веб-страниц и классы, работающие с API обрабатываемых сервисов. Для каждого из обрабатываемых сервисов написан отдельный парсер, работа которого определяется структурой исследуемой страницы. Для обеспечения безопасного доступа к данным и решения возможных проблем доступа используются такие инструменты, как:

- библиотеки, предоставляющие функции для удобной работы с API некоторых сервисов, например Apple Music API, VK API;
- Selenium – этот инструмент также позволяет выполнить скрипт и обратиться к HTML-элементу с помощью CSS-селектора;
- прокси.

*Структура сервиса
музыкальной аналитики
и рейтингования артистов на основе
их популярности в социальных сетях*

Сервис музыкальной аналитики и рейтингования артистов на основе их популярности в социальных сетях предоставляет доступ к данным в пользовательском интерфейсе и нескольких модулях API для прямого получения данных. Пользователи могут взаимодействовать с помощью нескольких приложений, которые позволяют:

- 1) войти в систему и извлечь данные об артистах с помощью ID артиста в системе FanStat;
- 2) извлечь данные об артистах с помощью ID в системе СберЗвук;
- 3) извлечь данные, полученные на основе моделей машинного обучения.

Структура сервиса представлена на рисунке.

Структурно сервис выстроен на основе интерфейсов прикладного программиро-

вания (API), что позволяет осуществлять последовательное взаимодействие с внутренними и внешними программными компонентами, а также возможности масштабирования за счет интеграции дополнительных функций. В рамках работы сервиса проработаны запросы POST (на предоставление данных в целом) и GET (получение данных из отдельного ресурса). Данные запросы относятся к категории HTTP-методов [5].

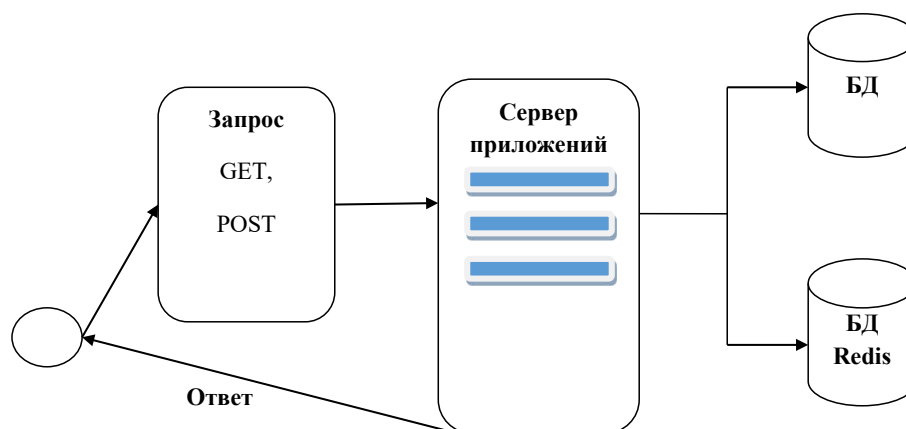
В рамках доступных для пользователя приложений возможно осуществление мониторинга артистов, рейтингования. В рамках дальнейшего масштабирования сервиса в качестве дополнительной аналитики можно провести мониторинг артиста по географическому критерию, на основе статистических данных по городам России.

Выборка артистов в обоих случаях проводится на основе имеющейся базы данных исполнителей из России и стран СНГ. Для поиска данных по артисту в базе данных используется его уникальный ID в конкретной системе, если это уместно, либо же его имя. Для сборки ID в некоторых системах также существуют отдельные парсеры.

Другая особенность сервиса – наличие двух систем управления базами данных (БД и Redis БД) для работы с основными и кэшированными данными.

Заключение

В основу проведенного исследования заложены более 20 источников данных и 86 рассматриваемых показателей для оценки популярности музыкальных артистов. Учет специфики и актуальности социальных сетей, радиостанций, стриминговых и поисковых сервисов, их распространенности среди пользователей в России позволяет поддерживать получаемые результаты в актуальной форме.



Структура работы сервиса

Результаты моделирования апробированы на 6000 артистов из России и стран СНГ. Полученные результаты агрегируются и представляются на разработанной платформе музыкальной аналитики FanStat.

Данный сервис позволяет увидеть и отслеживать динамику популярности, рейтинговать исполнителей, а также увидеть популярность артистов в региональном разрезе. Данная платформа может являться основой для проработки рекламных кампаний, определения музыкальных трендов, формирования музыкальных чартов на стриминговых сервисах и организации концертной деятельности.

Список литературы

1. Щепилова Г.Г., Круглова Л.А. Телеканалы и социальные сети: специфика взаимодействия // Вестник Московского университета. Серия 10: Журналистика. 2018. № 3.
2. Digital 2021 October Global Statshot Report [Электронный ресурс]. URL: <https://datareportal.com/reports/digital-2021-october-global-statshot> (дата обращения: 20.05.2022).
3. Аудитория социальных сетей и мессенджеров в 2021 году [Электронный ресурс]. URL: <https://blog.skillfactory.ru/> (дата обращения: 20.05.2022).
4. Завьялова Н.Б., Головина А.Н., Завьялов Д.В., Дьяконова Л.П., Мельников М.С., Сагинова О.В., Сагинов Ю.Л., Семенов А.В., Скоробогатых И.И., Строганов И.А. Методология и методы научных исследований в экономике и менеджменте: пособие для вузов. М.; Екатеринбург, 2014. 282 с.
5. Машнин Т.С. Технология Web-сервисов платформы Java. БХВ-Петербург, 2012. С. 115. 560 с.