

ТЕГИРОВАНИЕ КОНТЕНТА КАК ОСНОВА БАЗЫ ДАННЫХ СТРИМИНГОВОГО СЕРВИСА

CONTENT TAGGING IN DATA BASE OF STREAMING SERVICE

A. Yakovlev

Summary. The paper reviews some aspects of the data base forming for streaming audio services. The streaming service data base is based on a tagging complex with the system of tags grouped in several levels of personalization.

Keywords: Streaming services, tagging for data base.

Яковлев Антон Дмитриевич

Генеральный директор, ООО «Griffon-records»
misterantonio@yandex.ru

Аннотация. В статье рассматриваются особенности формирования базы данных стримингового сервиса на основе комплексного тегирования музыкального контента. База данных стримингового сервиса построена на комплексе тегирования, базирующимся на системе тегов, распределенных по нескольким уровням персонализации.

Ключевые слова: классификация, стриминговые сервисы, тегирование как основа базы данных.

Любая радиостанция стремится выстраивать формат своего вещания на основе музыкальных предпочтений определенной категории слушателей, так называемой *целевой аудитории* или *фокус-группы*, осуществляя отбор соответствующей музыки. На сегодняшний день современное FM-радиовещание достигло максимальных возможностей в части охвата слушательской аудитории и технических возможностей. При этом вне радиоэфира остаются направления музыки, неспособные стать основой радиовещания новых FM-радиостанции. Как правило, такие направления имеют недостаточное количество слушателей, для того, чтобы станция смогла эффективно монетизировать внимание соответствующей аудитории.

Вследствие этого обстоятельства, технологически продвинутые пользователи предпочитают *стриминговые сервисы*, такие как *Pandora, Spotify, Deezer* [1], [2], [3]; и другие; с одной стороны, имея возможность слушать только любимую музыку без оглядки на обязательные эфирные компромиссы, а с другой — получая регулярные обновления композиций, что делает услуги подобных сервисов разнообразными и комфортными для восприятия. Таким образом, в стриминговых сервисах центр форматного внимания сместился с интересов целевой группы слушателей на одного человека с его личными музыкальными предпочтениями. В этом случае единственный слушатель становится целевой аудиторией. В результате на смену десятков различных общепризнанных форматов приходит миллион неповторимых, уникальных форматов.

В основе базы данных любого стримингового сервиса находится *тегирование* — описание объекта (музыкальной композиции) с помощью ряда значений (*дескрипторов*), раскрывающих общие и индивидуальные характе-

ристики этого объекта. Тегирование должно учитывать разнообразную природу музыкальных композиций, специфику восприятия их человеком. Решение подобных масштабных задач всегда оказывается наукоемким, в частности, в тегировании для стримингового сервиса необходимы знания и практический опыт в психологии, лингвистическом анализе, информатике, музыкальном анализе.

1. Система тегирования музыкальных композиций

Автором разработана (реализована) система тегирования на основе комплексного подхода к анализируемому материалу. Приведем тезаурус основных определенных, используемых в рассматриваемой системе.

Значение (дескриптор) — единица тегирования, раскрывающая один из элементов содержания композиции. Последняя версия комплекса предоставляет 636 различных дескрипторов.

Атрибут — набор дескрипторов, объединенных в группу и описывающих одно из содержаний композиции. Один атрибут в базе данных сервиса тегирования содержит от 5 до 67 значений в зависимости от направленности музыки.

Блок — несколько атрибутов, описывающих результаты одного из методов анализа композиции. База данных сервиса тегирования содержит 3 блока: информационный, психологический и музыкальный. Каждый из этих блоков описывает самостоятельную область композиции.

Паспорт композиции — совокупность всех блоков, относящихся к одной композиции. Паспорт композиции

есть законченное описание индивидуально-выразительных средств композиции.

Шаблон исполнителя — совокупность паспортов композиций одного исполнителя, раскрывающая типичные и доминирующие дескрипторы его творчества.

Жанровый шаблон — совокупность шаблонов всех исполнителей, определяющая типичные и доминирующие характеристики всех исполнителей, относящихся к данному жанру.

Вокальный шаблон — часть шаблона исполнителя, раскрывающая дескрипторы вокального своеобразия исполнителя.

Дескрипторы разделены на 3 уровня персонализации:

Общий уровень = {качество звучания; темп; аранжировка; танцевальность; жанр}.

Детализирующий уровень = {национальный колорит; вокал; размер; использование по местоположению; плотность фактуры}.

Уточняющий уровень = {инструментальные партии; фактура; соло; диапазон}.

2. Блоки анализа

Ниже даны характеристики каждого из блоков анализа, применяемых в рассматриваемой системе тегирования.

2.1. Блок информационного анализа

Дескрипторы информационного блока, несомненно, несут важное значение, но не имеют принципиального значения в персонализации потока. Этот блок активно используется в традиционном радиовещании, но утрачивает доминирующее значение в стриминговых сервисах. Если название трека или группы для эфирной жизни не играет определяющей роли, то в стриминге — это возможность предложить пользователю, в случае его интереса, иные композиции данного коллектива или исполнителя.

В данный блок входят следующие дескрипторы: *название трека, группа/исполнитель, издатель, правообладатель, год издания, название альбома, аватар, текст композиции, разметка трека, оригинальность, жанр, контекст, похожие, тема.*

Метки

Расстановка меток корректирует время начала и конца воспроизведения трека, способствует плавности пе-

реходов между композициями в стриминговом потоке, отмечает временные границы звучания солиста, устанавливает точки входа и выхода, необходимые для автоматизации запуска рекламных аудиороликов, спотов, джинглов.

Контекст

Предполагает тематическую предрасположенность композиции к воспроизведению в определенном контексте: время (например, в канун Нового года) или локация (например, зал ожидания аэропорта или свадебный салон). Атрибут представлен в базе данных 44 активными значениями.

Тема/тематика

Для данного атрибута автор использовал определения, ориентируясь исключительно на собственные представления, которые корректировались в процессе увеличения объемов практического тегирования. По этой причине можно встретить такие определения как: *Любовь плюс, Любовь Минус, Просто любовь, За жизнь, Бунтарь, Документальный* и т.д.

Практическое тегирование не только сформировало устойчивую систему названий, поддержанную редакторами, но и обусловило трехуровневую систему тегирования поэтического образа:

1. Композиции по Содержанию рассматриваются на основе использования поэтических приемов аллегорий и метафор; для этих целей в медиабазе используется одно из 6 значений.
2. Атрибут **Тематика** компонуется в базе данных из 37 основных тем;
3. Каждая тема, в свою очередь, разделяется на подтемы, предоставляя в базу данных от 2 до 27 значений.

Подобная иерархия позволяет относить композиции как к широким тематическим группам, так и персонализировать их до узкотематических групп по интересам. Достаточно часто возникают сложности со строгой принадлежностью композиции к одной тематике, поэтому, в целях повышения объективности тегирования, уменьшения значения контекста и субъективности восприятия в работе редакторов при определении содержания и оценке тематики композиции, автором было введено понятие **тематического приоритета**, позволяющего редактору отмечать до трех тем одновременно в одной композиции с выставлением приоритетного значения между ними. Работа с разными музыкальными жанрами показывает доминирование в каждом из них определенного набора **атрибутов тем** и **значений тем**.

Время создания композиции (год издания)

Значения дескриптора выстроены по уменьшению временного периода, т.е. чем ближе к сегодняшнему дню, тем меньше временной шаг: *16–19 века; 1930–90 года; 2000–2016 года*.

Ненормативная лексика подразделяется на *Присутствие неформатной лексики; Присутствие неологизмов; Черный юмор*.

Дополнительная информация необходима для внесения пояснения редактором по имеющейся у него информации по треку.

Ошибки в базе представляют собой своего рода блокнот, позволяющий редактору, не отвлекаясь от поставленных главным редактором задач, сохранять проблемы, определяемые при прослушивании трека, например брак в звучании щелчки, провалы, необходимость замены трека и т.д.

2.2. Блок психологического анализа

Психологический анализ в основе тегирования не предоставляет столь развернутого инструментария для объективных оценок, как анализ музыкальный. Опираясь на понятия психических процессов восприятия, ментальных репрезентаций музыки, трудно получить объективные данные, способные релевантно группировать композиции по отобраным психологическим значениям. Сознательно отложив вопрос восприятия, формирования психологически схожих групп на следующий этап развития психологического тегирования, автор ограничился внесением в базу данных максимально удовлетворяющих требованиям объективности назначения тегов. К первоочередному ряду таких значений автор относит психологические явления, связанные с активностью и настроением.

2.3. Блок музыкального анализа

Ниже автор приводит характеристики некоторых основных дескрипторов блока музыкального анализа. Некоторые дескрипторы были намеренно вынесены за рамки данной статьи.

Жанр. Этот дескриптор является совокупностью всех музыкально-выразительных средств, используемых в комплексе тегирования, и представляет собой одну из точек входа для слушателя в стриминговый сервис. С учетом того, что подавляющее большинство слушателей предпочитает несколько жанров, только глубина тегирования предоставляет возможность стримингово-

му сервису найти точный путь к сочетанию музыкально-психологических средств, позволяющих формировать релевантное предложение пользователю.

Любой жанр при ближайшем рассмотрении не однороден, внутри него есть более или менее радикальные направления, отличающие конкретное произведение от основного жанра своим, хоть и родственным, набором выразительных средств. Задача жанровых дескрипторов в конечном итоге максимально точно определять этот комплекс выразительных средств, переведенных в наш случай в комбинации дескрипторов. Поэтому жанровое строение базы данных предполагает наличие *основного жанра* композиции и *поджанров*, детализирующих основной жанр.

Например, основной жанр *Рок*, детализируется следующим набором поджанров: *Русский рок, Рокопопс, Тяжелый рок, Джаз-рок, Альтернативный рок, Панк-рок, Фолк-рок, Глэм-рок, Софт-рок, Электро-рок, Дрим-рок, Инди-рок, Психоделический рок, Британский рок, Эстрадный рок, Гранж, Блюз-рок, Готик рок и т.д.*

Помимо общеупотребительных жанровых терминов в медиа базе появляются и новые, редко используемые (например, *блатная электронная, рэп-подземка* и т.д.), но дающие четкое понимание причинно-следственных связей.

Дескриптор **Вокал** состоит из полей *Пол; Диапазон/Ансамбль; Тембр* (определяет основные категории постановки голоса и манер звукоизвлечения); *Владение голосом* (ранжирует исполнительское мастерство с технической точки зрения от самодеятельности до свободного владения); *Вибрация* (характеризует наличие/отсутствие вибрато в голосе и его характеристику); *Вокальная подача* (определяет вокальную активность конкретного произведения, в случае доминирующей подачи отмечает именно его); *Голосовые особенности* (голосовую специфика маркирующая грассирование, придыхание, пение в нос, хрипотцу и т.д. Всего в общей сложности включает в себя 43 значения, что в совокупности с выше изложенными предоставляет отличный редакторский инструмент для детального описания одного из важнейших выразительных элементов песенного жанра — вокала).

Инструментальное **соло** относится к выразительным элементам первого уровня. В песенном творчестве, как правило, солирующий инструмент появляется в качестве замены вокальной партии.

Включает в себя следующие поля: *гитарный блок акустический, гитарный блок электронный, сольные инструменты симфонического оркестра; группы сим-*

фонического оркестра; вокал; народные инструменты; синтезированные инструменты.

Качество звучания подразделяется по следующим характеристикам:

1. Обобщенный тембр композиции = {Живой звук | Акустические инструменты | Электронный}.
2. Эпоха = {Старое | Современное}.
3. Фактурная насыщенность = {Прозрачная | Фоновая | Плотная | Разнородная}.

3. Автоматизация тегирования

Результаты исследований автора в области тегирования в настоящее время позволяют говорить о возможности применения автоматизации тегирования в следующих характеристиках: *темп; качество звучания; разметка трека; фактура*.

Автоматизация определения темпа композиции.

Темп является одним из основных компонентов восприятия музыки. Темп делегирует в базу данных 6 темповых значений: *Очень медленный/Медленный/Средне-медленный/ Средний/Средне-быстрый/Быстрый*. В некоторых случаях именно темп может выступать в качестве доминирующего дескриптора в формировании персонального потока. В частности, в плейлистах с условно спокойной музыкой он является доминирующим тогда, когда тематика или национальный колорит могут отсутствовать вообще. В электронной музыке темповый коридор вообще является обязательным условием принадлежности трека к одному из поджанров.

При всей условности и субъективности определения темпа, это один из атрибутов тегирования, сравнительно легко поддающийся автоматизированному анализу. Задача автоматизации значения темпа в танцевальной музыке не представляется проблемной: существует достаточно много сервисов, обеспечивающих довольно точные результаты. Сложности начинаются в музыке, отличной от клубно-танцевальных жанров, например, в романсах или иных «свободных» темпах. Ключевым вопросом решения автоматизации темпа является задача метроритма и внутреннего движения. Автоматизация темпа позволит точно выставить индивидуальную темповую партитуру для любого пользователя.

Дескриптор **качество звучания** позволяет установить принадлежность музыкальной композиции к тому или иному периоду (*старое | музыка 60–70х | музыка 80х | музыка 90х и современная музыка*).

Автоматизация достигается за счет использования инструментария технических средств, таких как спектральный анализ и временной анализ. Известная тенденция продюсеров увеличить громкость произведения позволяет сравнительно легко отделить музыку послевоенных лет от современной клубной музыки. Для увеличения точности делимости используется спектральный анализ, в частности параметр **крутизна спектра**.

Таким образом, вывод данного инструмента можно представить множеством точек на плоскости. При этом каждому периоду будет соответствовать своя область рассматриваемой плоскости. За классификацию отвечает нейронная сеть, прошедшая обучения на выборке из 500 музыкальных композиций различных жанров и относящихся к различным периодам.

4. Система пользовательских действий

Музыка в радиопотоке существует по вполне определенным статистическим законам: некоторые композиции являются ночными, некоторые — преимущественно дневными; некоторые треки реципиенты предпочитают слушать в течение рабочей недели, другие, наоборот, по выходным. В стриминговом сервисе, в отличие от радионного, возникает интерактивная возможность воздействия на музыкальный контент. Поэтому первоначально заложенный статус композиций в результате пользовательского выбора трансформируется в персональный выбор отдельного слушателя. Именно поэтому пользовательское действие является важным инструментом подстройки потока, а мотивация пользователя к совершению пользовательского действия является важной задачей в решении персонализации потока.

Осуществление пользовательского действия переводит композицию с общего статуса в персонализированные. К основным пользовательским действиям отнесены установка «лайка», установка «дизлайка», «перемотка» (переход к следующей композиции в потоке).

Пользовательское действие является довольно специфической характеристикой. Оно сильно зависит от конкретного пользователя. Реципиент А осуществляет перемотку по причине нежелания слушать трек в данную минуту времени, в то время как для реципиента Б это действие равносильно дизлайку.

Заключение

Данная работа является результатом практической деятельности автора, как профессионального композитора, много лет руководившего различными популярными музыкальными радиостанциями и практического

тегирования более полутора десятка тысяч разножанровых композиций. Задача данной системы — не подбор любимых песен, а формирование набора музыкальных рекомендаций на основании предпочтений пользователя. При понимании системой того, какие признаки доминируют при выборе слушателем тех или иных композиций, рекомендательная система предлагает слушателю новые, с высокой вероятностью того, что данные композиции будут приняты пользователем благосклонно.

Проблема рассматриваемых сервисов заключается в том, что пользователю невозможно предложить заведомо релевантную выдачу, то есть, проблема так называемых холодных стартов является общей для всех рекомендательных систем. С момента обращения к рекомендательной системе до получения необходимого пользователю результата должно пройти несколько действий пользователя. На каждом из них задействованы определенные дескрипторы, расположенные от общего к частному они, и представляют собой алгоритм персонализации.

ЛИТЕРАТУРА

1. Corinne Loiacono. Internet Radio: An Analysis of Pandora and Spotify. Senior Capstone Project, Spring 2014. http://digitalcommons.bryant.edu/cgi/viewcontent.cgi?article=1013&context=honors_mathematics (Дата обращения 20.11.2016)
2. Официальный сайт сервиса Pandora <http://www.pandora.com/about> (Дата обращения 20.11.2016)
3. Официальный сайт сервиса Spotify <http://www.spotify.com> (Дата обращения 20.11.2016)

© Яковлев Антон Дмитриевич (misterantonio@yandex.ru)

Журнал «Современная наука: актуальные проблемы теории и практики»