

ОСОБЕННОСТИ ПОСТРОЕНИЯ АРХИТЕКТУРЫ ИНФОРМАЦИОННОЙ СИСТЕМЫ, ИСПОЛЬЗУЮЩЕЙ АЛГОРИТМЫ МАШИННОГО СЛУХА

FEATURES OF THE DESIGN OF ARCHITECTURE OF INFORMATION SYSTEMS USING MACHINE HEARING ALGORITHMS

*T. Repetskaya
N. Repetskaya*

Summary. Features of the design of architecture of information systems using machine hearing algorithms are considered. The classification of typical tasks solved using machine hearing algorithms is carried out, taking into account the requirements and limitations of the modern IT infrastructure. As a result of this research, an approach to the design of the architecture of software solutions was developed taking into account the specifics of such algorithms.

Keywords: machine hearing, architecture of information systems, design of the architecture of software solutions.

Репецкая Татьяна Олеговна

Национальный исследовательский университет
«Высшая школа экономики» (г. Москва)
torepetskaya@edu.hse.ru

Репецкая Наталья Викторовна

К.э.н., доцент, ИАТЭ НИЯУ МИФИ (г. Обнинск)
nataliavr2005@ya.ru

Аннотация. Рассмотрены особенности построения архитектуры информационной системы, использующей алгоритмы машинного слуха. Проведена классификация типовых задач, решаемых с помощью алгоритмов машинного слуха, с учетом требований и ограничений со стороны современной ИТ-инфраструктуры. В результате проведенного исследования сформирован подход к разработке архитектуры программных решений с учетом специфики таких алгоритмов.

Ключевые слова: машинный слух, архитектура информационной системы, разработка архитектуры программных решений.

Технологии искусственного интеллекта, перспективного научного направления, ставшего популярным за последние 10 лет, развиваются все с большими темпами и предоставляют пользователям компьютерной техники разнообразный инструментальный функционал. Системы искусственного интеллекта помогают людям решать задачи разного уровня сложности, как в научно-производственной сфере, так и в бытовой, применяются в системах безопасности, биометрических системах, технологиях «умного дома» и «умного города».

Одним из основных понятий искусственного интеллекта является машинное восприятие (Machine Perception), к которому относятся машинное зрение (Machine Vision), машинный слух (Machine Hearing) и машинное осязание (Machine Touch) [4].

Наиболее распространенными и более изученными из систем машинного восприятия сегодня являются системы машинного зрения. Они позволяют компьютеру «видеть» объекты окружающего мира, различать их и на основе анализа данных предлагать наиболее оптимальные решения поставленных задач. Немало важным также является возможность научить компьютер «слышать» и идентифицировать различные звуки окружающего мира, что объясняет стремительное развитие

и рост популярности систем такого типа, например, систем распознавания речи.

Благодаря достижениям в области информационных технологий, алгоритмы машинного слуха делают возможным решение множества задач в сфере банковской деятельности, медицины, безопасности, робототехники и др., что делает это направление предметом интереса как науки, так и государства, и бизнеса [3]. Однако степень изученности области машинного слуха по-прежнему остается низкой в сравнении с другими областями машинного восприятия.

Машинный слух — это способность компьютерной системы с помощью аппаратных и программных средств воспринимать, обрабатывать и анализировать данные в виде звуковых волн [3]. Выделяют следующие направления машинного слуха: распознавание и анализ музыки [7], распознавание речи [5] и распознавание звуков окружающей среды [1, 4], в каждом из которых можно выделить свойственные им типы решаемых задач (рис. 1).

Каждое из существующих направлений машинного слуха имеет свою специфику, учитывающую требования к входным данным, объем обрабатываемых данных, нагрузку на систему, инструменты обработки и другие критерии. Поэтому очень важно развивать понимание



Рис. 1. Направления машинного слуха (типы решаемых задач, примеры существующих систем)

| | Распознавание и анализ музыки (Music Recognition) | | | Распознавание речи (Speech Recognition) | | | Распознавание окружающей среды (ESR) | |
|------------------------------------|---|---|---|---|---|--|--------------------------------------|--------------------------------|
| | Распознавание трека | Поиск аналогичных треков | Распознавание верхнеуровневых признаков трека | Распознавание слов | Распознавание слитной речи | Идентификация паттернов речи | Определение звуковых сцен | Распознавание звуковых событий |
| Требования к входным данным | Низкий уровень шума | Отсутствие шума/Низкий уровень шума | Отсутствие шума/Низкий уровень шума | Низкий уровень шума | Низкий уровень шума | Низкий уровень шума | Не имеет значения | Не имеет значения |
| Обработка входных данных | В режиме реального времени/Автономно | В режиме реального времени/Автономно | В режиме реального времени/Автономно | В режиме реального времени | В режиме реального времени | В режиме реального времени | В режиме реального времени | В режиме реального времени |
| Объем обрабатываемых данных | Средний | Очень большой | Средний | Средний | Большой | Средний | Очень большой | Очень большой |
| Инструмент обработки | Метод сигнатур | Метод акустических отпечатков/Машинное обучение | Метод обертонового ряда | Скрытая Марковская модель | Скрытые Марковские модели/Машинное обучение | Скрытые Марковские модели/Нейронные сети | Машинное обучение | Машинное обучение |
| Объем хранимых данных | Очень большой | Очень большой | Низкий | Средний | Очень большой | Средний | Очень большой | Очень большой |
| Нагрузка на систему | Средняя | Средняя/Высокая | Средняя | Средняя | Средняя/Высокая | Средняя | Высокая | Высокая |
| Надежность | Высокая | Высокая | Средняя | Высокая | Высокая | Высокая | Высокая | Высокая |

Рис. 2. Классификация типовых задач машинного слуха и их специфика

тех проблем, которые могут быть решены с помощью алгоритмов машинного слуха, а также тех ограничений и требований, которые влечет за собой использование этих алгоритмов в контексте современной ИТ-инфраструктуры [2, 6].

В рамках основных направлений машинного слуха можно выделить более узкие задачи, решаемые с помощью алгоритмов машинного слуха. Изучив научные статьи зарубежных авторов и проанализировав различные сферы человеческой деятельности, в которых потенци-



Рис. 3. Взаимовлияние технологий машинного слуха и ИТ-инфраструктуры (на примере распознавания речи)

ально могут быть использованы технологии машинного слуха, выделим основные типовые задачи и проведем их классификацию с учетом требований и ограничений (рис. 2).

Развитие информационных технологий машинного слуха, зарождавшихся еще в середине XX века (Audrey, Shoebbox, Harpy, Dragon, GoogleVoiceSearch, Siri) оказывало влияние на развитие ИТ-инфраструктуры используемых их информационных систем (рис. 3).

А развитие ИТ-инфраструктуры (появление мощных процессоров, хранилищ данных, облачных вычислений), в свою очередь, существенно повлияло на эволюцию технологий распознавания речи и машинного слуха в целом.

Такое взаимовлияние информационных технологий машинного слуха и ИТ-инфраструктуры позволяет сделать вывод, что архитектура информационной системы, использующей алгоритмы машинного слуха, зависит от класса решаемых задач и должна учитывать ограничения и требования, предъявляемые со стороны современной ИТ-инфраструктуры к каждому классу задач. Поэтому при разработке архитектуры ИС таких систем важно понимать, какие задачи машинного слуха будут решаться, и правильно определять требования к разработке.

На основе свода знаний системной инженерии SEBoK [6], методологии построения корпоративных информационных систем Фаулера [8] и с учетом специфики задач машинного слуха сформируем подход к разработке архитектуры информационных систем, использующих алгоритмы машинного слуха. Процесс разработки архитектуры информационных систем, использующих алгоритмы машинного слуха, включает семь основных этапов. Рассмотрим их, и продемонстрируем подход к разработке архитектуры подобной ИС на практиче-

ском примере разработки архитектуры информационной системы сервиса по подбору музыки под настроение.

Описание концепта информационной системы (ИС)

На первом этапе разработки архитектуры информационной системы необходимо описать концепцию ИС в общем виде, отобразить функциональные требования к ИС и пользовательские сценарии для дальнейшей группировки требований.

Например, рекомендательный веб-сервис, который анализирует музыку и формирует плейлисты под настроение пользователя, основываясь на структурной составляющей композиции (ритм, скорость, тональность и пр.), и общепринятых характеристиках (жанр, гармония и др.).

Определение назначения и задач ИС

На следующем этапе разработки архитектуры информационной системы необходимо в рамках описанной концепции сформулировать предназначение разрабатываемой ИС и решаемые системой задачи.

Важно также показать, какие преимущества дает разрабатываемая система пользователям. Целью этого этапа является определение функционала системы на уровне бизнес-требований и их приоритетов.

Например, назначение рекомендательного веб-сервиса — формирование плейлистов песен, хранящихся в памяти системы, за счет формирования предсказаний текущего настроения пользователя в режиме реального времени. Такой функционал предполагает работу системы в режиме реального времени, хранение большого объема данных и одновременную доступность большому количеству пользователей.

| Функциональная архитектура | Временная архитектура | Поведенческая архитектура |
|--|--|--|
| Выделение физических параметров трека | Каждый раз при загрузке новой композиции в систему | <ul style="list-style-type: none"> • Кластеризация базы треков • Формирование профилей пользователей (информация о любимых/не любимых композициях). • Выбор одного из кластеров для поиска наилучшей песни (песен). • Выбор следующей песни (песен) для рекомендации из кластера |
| Классификация треков, содержащихся в базе данных сервиса | Ежемесячно при условии добавления новых композиций в базу данных в течение периода | |
| Определение паттернов поведения пользователя и формирование кластеров его настроений | Регулярно при каждом использовании | |
| Формирование плейлиста | Регулярно при каждом использовании | |
| Идентификация пользователя | Регистрация – единожды, авторизация – регулярно | |

Рис. 4. Логическая архитектура сервиса по подбору музыки под настроение

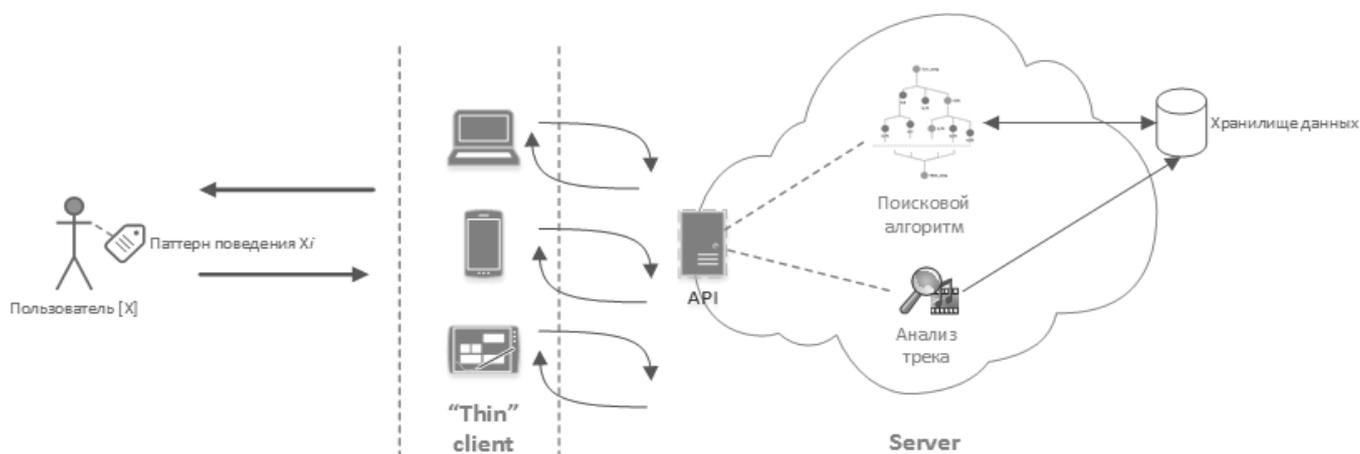


Рис. 5. Компоненты ИС сервиса по подбору музыки под настроение

Классификация решаемой задачи

На основе описанной задачи и бизнес-требований с учетом их приоритизации (этап 1), классифицировать решаемую задачу согласно сформированной классификации типовых задач (таблица 1) с целью заключительного формирования бизнес-требований к системе.

Пример: две типовые задачи направления Music Recognition: распознавание верхнеуровневых признаков трека и поиск аналогичного трека.

Бизнес-требования к информационной системе:

- ♦ работа в режиме реального времени;
- ♦ возможность хранения и обработки большого объема данных одновременно;
- ♦ большое количество пользователей сервиса;

| ТРЕБОВАНИЕ (ЭТАП 3) | “ТОЛСТЫЙ” КЛИЕНТ | “ТОНКИЙ” КЛИЕНТ | СМЕШАННЫЙ КЛИЕНТ |
|--|--|--------------------|------------------|
| Система должна работать в режиме реального времени | Да, безотносительно других пользователей | Да | Да |
| Должно быть обеспечено хранение и обработка большого количества данных | Нет | Да, но не локально | Да |
| | | | |

Рис. 6. Соотнесение ограничений и требований задачи с моделями клиент-серверного взаимодействия

| | Вес параметра | “Толстый” клиент | “Тонкий” клиент | Смешанный клиент |
|-------------------------------|---------------|------------------|-----------------|------------------|
| Быстрота реагирования | 0,15 | 3 | 2 | 2 |
| Время задержки | 0,13 | 3 | 1 | 2 |
| Пропускная способность | 0,18 | 1 | 3 | 3 |
| Чувствительность к загрузке | 0,14 | 3 | 2 | 2 |
| Эффективность | 0,12 | 2 | 2 | 3 |
| Мощность | 0,20 | 1 | 3 | 3 |
| Способность к масштабированию | 0,08 | 1 | 3 | 2 |
| | | 1,96 | 2,33 | 2,50 |

Рис. 7. Выбор архитектурного решения с помощью метода балльной оценки

- ♦ содержание минимума/отсутствие шума в треках;
- ♦ выдерживать высокую нагрузку, быть высоконадежной;
- ♦ использование алгоритмов машинного слуха (метода акустических отпечатков, обертонового ряда, машинное обучение).

Формирование логической архитектуры

Определение бизнес-требований к системе позволяет на следующем этапе выстроить логическую архитектуру ИС с использованием минимально необходимых

групп архитектур: функциональной, поведенческой и временной. Целью этого этапа является соотнесение бизнес-требований и требований пользователей на уровне функциональных требований.

Логическая архитектура для рассматриваемого примера представлена на рисунке 4.

Выделение компонентов ИС

На следующем этапе разработки архитектуры информационной системы необходимо на основе выстроенной логической архитектуры выделить компоненты

системы и специфические инструменты, необходимые для их реализации с целью извлечения системных требований к программному решению (рис. 5).

Соотнесение ограничений задачи и требований к программному решению с требованиями к ИТ-инфраструктуре

Главной целью этого этапа является формирование вариантов построения архитектуры ИТ-инфраструктуры программного решения. Для этого необходимо группы требований к ИС, определенных на третьем этапе с учетом классификации типовых задач наложить на каждую из возможных моделей клиент-серверного взаимодействия с учетом их специфики. Пример соотнесения ограничений задачи и требований к программному решению с требованиями к ИТ-инфраструктуре для сервиса по подбору музыки под настроение представлен на рисунке 6.

Оценка вариантов и выбор архитектуры ИС

На заключительном этапе разработки архитектуры ИС необходимо выбрать один из сформированных на предыдущих этапах вариант построения архитектуры ИТ-инфраструктуры. Для принятия решений при выборе программного обеспечения целесообразно применять такие методы оценки, как балльная оценка, экспертно-балльная оценка и др.

При осуществлении оценки необходимо учитывать следующие параметры, характеризующие функционирование информационной системы согласно методологии Фаулера [8]:

- ◆ быстроту реагирования,
- ◆ время задержки,
- ◆ пропускную способность (производительность) системы,
- ◆ чувствительность к загрузке (зависимость времени отклика от загрузки),
- ◆ эффективность,
- ◆ мощность,
- ◆ способность к масштабированию.

В рассматриваемом примере разработки архитектуры ИС, использующей алгоритмы машинного слуха, выбор архитектурного решения выполним с помощью метода балльной оценки (рис. 7).

В результате проведенной оценки выбран вариант построения архитектуры информационной системы сервиса по подбору музыки под настроение по смешанной модели клиент-серверного взаимодействия.

Таким образом, сформирован и показан на практике подход к построению архитектуры информационной системы, использующей алгоритмы машинного слуха, который позволяет учесть специфику решаемых с помощью таких алгоритмов задач и получить на выходе комплексный набор требований. В результате применения представленного подхода формируется набор возможных решений поставленной задачи, по которым экспертная группа проектной команды разработки осуществляет окончательный выбор архитектурного решения. Сформированный подход к разработке архитектуры информационной системы, использующей алгоритмы машинного слуха, позволит более правильно определить исходные требования к разработке, минимизировать возможные ошибки разработки и, тем самым, сократить затраты на разработку ИС.

ЛИТЕРАТУРА

1. Aytar Y., Vondrick C., Torralba A. Soundnet: Learning sound representations from unlabeled video // Advances in Neural Information Processing Systems. — 2016. — С. 892–900.
2. Laan S. IT Infrastructure Architecture-Infrastructure Building Blocks and Concepts Third Edition. — Lulu.com, 2017.
3. Lyon R. F. Machine hearing: An emerging field [exploratory dsp] // IEEE signal processing magazine. — 2010. — Т. 27. — № . 5. — С. 131–139.
4. Piczak K. J. Environmental sound classification with convolutional neural networks // Machine Learning for Signal Processing (MLSP), 2015 IEEE 25th International Workshop on. — IEEE, 2015. — С. 1–6.
5. Pinola M. Speech recognition through the decades: How we ended up with siri // Web log post. TechHive. IDGTechNetwork. — 2011. — Т. 2.
6. Pyster A. et al. Guide to the systems engineering body of knowledge (SEBoK) v. 1.0. 1 // Guide to the Systems Engineering Body of Knowledge (SEBoK). — 2012.
7. Wang A. The Shazam music recognition service // Communications of the ACM. — 2006. — Т. 49. — № . 8. — С. 44–48.
8. Фаулер М. Архитектура корпоративных программных приложений: пер. с англ / М.: Издательский дом Вильямс. — 2006. — 544 с.

© Репецкая Татьяна Олеговна (torepetskaya@edu.hse.ru), Репецкая Наталья Викторовна (nataliavr2005@ya.ru).

Журнал «Современная наука: актуальные проблемы теории и практики»