УДК 004.896:621.3.049.77

ИСПОЛЬЗОВАНИЕ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА ПРИ ФОРМИРОВАНИИ БАЗЫ ДАННЫХ IMEI-КОДОВ МОБИЛЬНЫХ УСТРОЙСТВ

Давронбеков Д.А., Исроилов Ж.Д.

Ташкентский университет информационных технологий имени Мухаммада ал-Хоразмий, Ташкент, Узбекистана, jamshid.isroilov@gmail.com

Аннотация: в работе рассматривается подход к формированию базы данных IMEI-кодов мобильных устройств с использованием методов искусственного интеллекта. В отличие от традиционного формирования реестров, основанного преимущественно на структурной и справочной проверке записей, предлагается дополнительно учитывать взаимосвязь IMEI с атрибутами мобильного оборудования и параметрами зарегистрированных событий. Сформирована структура входных данных, включающая идентификатор устройства, TAC, сведения о производителе и модели, временные характеристики и параметры регистрации. Предложена последовательность предварительной обработки данных, формирования признакового пространства и интеллектуальной оценки качества записей. Практическая значимость подхода заключается в возможности автоматизации выявления неполных, противоречивых и аномальных записей при формировании централизованной базы данных.

Ключевые слова: IMEI, база данных, искусственный интеллект, TAC, мобильное устройство, качество данных, аномалия, телекоммуникационная сеть

I. ВВЕДЕНИЕ

IMEI (International Mobile Equipment Identity) является идентификатором мобильного оборудования, используемым мобильными сетями для идентификации устройства. Согласно материалам GSMA, IMEI представляет собой 15-значный идентификатор, а TAC используется для идентификации типа, производителя и модели мобильного устройства. GSMA также поддерживает глобальные сервисы, связанные с распределением TAC, сведениями об устройствах и статусом идентификаторов [1].

В современных мобильных сетях количество устройств и объём формируемых служебных данных постоянно увеличиваются. В связи с этим база данных IMEI должна рассматриваться не только как совокупность идентификаторов, но и как информационная система, обеспечивающая хранение, проверку и актуализацию взаимосвязанных сведений о мобильном оборудовании.

Традиционный подход к формированию базы данных предполагает выполнение детерминированных операций: проверку длины IMEI, контроль контрольной цифры, определение TAC и сопоставление идентификатора со справочными сведениями о производителе и модели. Однако такие проверки не позволяют в полной мере оценить качество записи в случае наличия неполных, противоречивых или нетипичных комбинаций атрибутов.

В этой связи особое значение приобретает интеллектуальный анализ совокупности характеристик IMEI-записей, позволяющий учитывать не только их формальную корректность, но и структурные, справочные, статистические и временные взаимосвязи. Применение методов искусственного интеллекта в данном случае позволяет перейти от отдельных процедур проверки к комплексной оценке качества записей и автоматизированному выявлению отклонений от установленных характеристик информационной базы [2].

Предлагаемый подход ориентирован на формирование концептуальной модели базы данных IMEI-кодов, в которой процедуры предварительной обработки и валидации дополняются интеллектуальным анализом признакового пространства. Такая организация обработки обеспечивает классификацию записей по степени соответствия заданным требованиям, выявление потенциально аномальных данных и повышение достоверности формируемой базы данных.

Базовой единицей предлагаемой информационной модели является запись о мобильном устройстве:

$$D_i=\{I_i,T_i,M_i,R_i,E_i,Q_i\}, \quad (1)$$

где I_i - IMEI устройства; T_i - TAC; M - сведения о производителе и модели; R_i - регистрационные параметры; E_i- временные и сетевые характеристики зарегистрированных событий; Q_i- показатели качества записи [1-7].

Структура IMEI может быть представлена следующим образом: *IMEI=TAC‖SNR‖CD,* где *TAC* - Type Allocation Code, *SNR* - серийная часть идентификатора, *CD* - контрольная цифра, а символ ‖обозначает конкатенацию полей.

GSMA указывает, что TAC имеет длину 8 цифр, после него следуют 6 цифр серийного номера, а последняя цифра IMEI используется как контрольная цифра.

Для каждой записи базы данных целесообразно хранить не только сам IMEI, но и дополнительные атрибуты:

$$X_i=\{TAC_i,SNR_i,V_i,M_i,C_i,t_i,L_i,S_i\}, \quad (2)$$

где *Vi* - версия программного обеспечения устройства, C_i - класс устройства, t_i - временная характеристика, L_i - информация о зоне регистрации, S_i - параметры источника данных.



Рисунок 1. Общая архитектура использования искусственного интеллекта при формировании базы данных IMEI-кодов мобильных устройства

РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

Такое представление позволяет перейти от простой таблицы идентификаторов к структурированной модели данных.

ПРИМЕНЕНИЕ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

Предлагаемый подход состоит из четырёх последовательных этапов:

1. получение и структурирование исходных данных;
2. детерминированная предварительная проверка;
3. формирование признакового пространства;
4. интеллектуальная оценка качества записи.

На первом этапе из различных источников формируется массив записей

$$D=D=\{D_1,D_2,\ldots,D_N\}. \quad (3)$$

В качестве признакового пространства используются структурные, справочные, информационные, статистические и временные характеристики записей базы данных IMEI. Они отражают корректность структуры идентификатора, соответствие IMEI и TAC справочным сведениям, полноту и согласованность атрибутов мобильного устройства, а также характеристики частоты и динамики регистрационных событий. Совокупность указанных признаков позволяет комплексно оценивать качество информационных записей и формировать входные данные для интеллектуальной модели классификации и выявления аномальных записей [3,5].

Для оценки качества записи может использоваться интеллектуальная модель:

$$S_i=f_\theta(F_i), \quad (4)$$

где f_θ - обученная модель искусственного интеллекта, а S_i - итоговая оценка записи.

Для практической реализации могут рассматриваться методы классификации, кластеризации или обнаружения аномалий. Выбор конкретного алгоритма должен определяться доступностью размеченной выборки. При наличии классов «корректная запись» и «некорректная запись» возможно применение контролируемого обучения. При отсутствии достоверной разметки более рациональным является использование методов unsupervised anomaly detection.

АЛГОРИТМ ИНТЕЛЛЕКТУАЛЬНОГО ФОРМИРОВАНИЯ БАЗЫ

Предлагаемый алгоритм можно представить следующей последовательностью:

$$D_i \rightarrow Preprocessing \rightarrow Validation \rightarrow Feature\ Extraction \rightarrow AI \rightarrow S_i \rightarrow Database.$$

На этапе предварительной обработки выполняются удаление дубликатов записей, нормализация полей, обработка пропущенных значений и приведение данных к единому формату.

После этого осуществляется детерминированная проверка IMEI. Записи, не соответствующие обязательным структурным требованиям, направляются на дополнительную проверку.

Важным является то, что AI-модель в предлагаемой архитектуре не заменяет нормативную проверку IMEI. Она выполняет дополнительную функцию интеллектуального контроля качества данных.

Таким образом, логика обработки имеет два уровня:

IMEI→формальная проверка→интеллектуальная проверка→решение о качестве записи.

Такое разделение позволяет избежать ситуации, при которой статистическая модель самостоятельно определяет нормативную корректность идентификатора. Предлагаемый подход изменяет принцип формирования базы данных IMEI с преимущественно статического хранения идентификаторов на интеллектуальный контроль качества информационных записей.

Особенностью модели является совместное использование детерминированных и интеллектуальных методов. Детерминированный уровень обеспечивает соблюдение формальных требований к идентификатору, тогда как AI-уровень анализирует многомерные зависимости между атрибутами записи.

Таблица 1. Структура полей и признаков базы данных IMEI-кодов мобильных устройств

Группа	Примеры данных
Идентификация	IMEI, TAC, SNR
Устройство	производитель, модель, тип
Программное обеспечение	версия ПО
Регистрация	дата, время, параметры события
Источник	источник формирования записи
Контроль	результаты формальной проверки
AI-анализ	признаки, оценка S_i класс
Состояние	подтверждённая, требующая проверки, аномальная

Такое представление согласуется с общей логикой современных систем управления идентификаторами устройств. GSMA Device Registry, например, предназначен для ведения сведений о TAC и глобального списка идентификаторов устройств, имеющих определённый статус.

При этом разработанный в работе подход следует рассматривать как надстроечный аналитический механизм, а не как замену существующих процедур управления IMEI [6,7].

ЗАКЛЮЧЕНИЕ

В работе рассмотрена задача формирования базы данных IMEI-кодов мобильных устройств с применением методов искусственного интеллекта. Показано, что традиционная проверка структуры IMEI является необходимой, но недостаточной для комплексной оценки качества информационной записи.

Предложена структурная модель записи, включающая идентификатор устройства, TAC, сведения о производителе и модели, регистрационные параметры и признаки качества данных. Разработана

последовательность обработки, объединяющая предварительную обработку, формальную валидацию, формирование признакового пространства и интеллектуальную классификацию.

Основным результатом является модель двухуровневого контроля, в которой нормативные проверки выполняются детерминированными процедурами, а методы искусственного интеллекта используются для выявления неполных, противоречивых и статистически нетипичных записей.

Дальнейшее исследование целесообразно направить на формирование экспериментальной выборки реальных регистрационных данных, сравнительную оценку алгоритмов Random Forest, Gradient Boosting, Isolation Forest и нейросетевых методов, а также определение показателей Precision, Recall, F1-score и ROC-AUC для количественной оценки эффективности предлагаемого подхода.

БЛАГОДАРНОСТЬ

Авторы выражают благодарность специалистам в области мобильных телекоммуникационных систем за консультации при формировании структуры исследуемых данных и постановке задачи контроля качества базы данных IMEI-кодов.

ЛИТЕРАТУРА

[1] Eoghan Casey. Handbook of Digital Forensics and Investigation. – Copyright Eoghan Casey. Published by Elsevier Inc. - 2010 – 559 p.

[2] Raj, Tilak. (2025). IMEI: An Essential Resource for Device Identification and Security. International Journal for Research in Applied Science and Engineering Technology. 13. 1368-1370. 10.22214/ijraset.2025.67560.

[3] М.С.Лохвицкий, Н.С.Мардер. Сотовая связь: от поколения к поколению. - М. Издательство ИКАР. - 2014 - 236 с.

[4] Д.Давронбеков, А.Кириакиди, Д.Елькин, Ж.Исроилов, М.Нурматова. Состояние и анализ рынка мобильных устройств для создания единой базы идентификации мобильных устройств по IMEI // International Scientific Journal«Science and World». - 2018. - №11(63). Vol.I. - P. 41-42.

[5]D. A. Davronbekov, J. D. Isroilov and B. I. Akhmedov, "Principle Of Organizing Database Identification On Mobile Devices By IMEI," 2019 International Conference on Information Science and Communications Technologies (ICISCT),Tashkent, Uzbekistan, 2019, pp. 1-5, doi: 10.1109/ICISCT47635.2019.9012000.

[6] Д.Давронбеков, К.Абдурахманов, З.Хакимов, Ж.Исроилов, А.Кириакиди, Д.Елькин. Некоторые вопросы регистрации и учета мобильных устройств по IMEI // International Scientific Journal «Science and World». -2019.-№8(72)-P.27-29

[7] Давронбеков Д., Исроилов Ж., Архитектура организации регистрации IMEI-кодов в базе данных информационной системы идентификации мобильных устройств //Наука и мир. - 2020. -№.10.-С.39-43.

USING ARTIFICIAL INTELLIGENCE TO BUILD A MOBILE DEVICE IMEI DATABASE

D. Davronbekov, J. Isroilov

Muhammad al-Khwarizmi Tashkent University of Information Technologies, Tashkent, Uzbekistan,
jamshid.isroilov@gmail.com

Abstract: this paper examines an approach to building a mobile device IMEI database using artificial intelligence methods. Unlike traditional registry creation, which relies primarily on structural and reference-based record validation, this approach proposes also considering the relationships between the IMEI and mobile equipment attributes as well as registered event parameters. An input data structure has been defined, comprising the device identifier, TAC, manufacturer and model details, timing characteristics, and registration parameters. A sequence is proposed for data preprocessing, feature space construction, and intelligent record quality assessment. The practical significance of this approach lies in the ability to automate the detection of incomplete, inconsistent, and anomalous records during the creation of a centralized database.

Keywords: IMEI, database, artificial intelligence, TAC, mobile device, data quality, anomaly, telecommunications network.