Conference of Academies for Applied Studies in Serbia (CAASS)
University Business Academy in Novi Sad, Faculty of Contemporary Arts, Belgrade

BOOK OF PROCEEDINGS

Editor: Dr Branko Savić, professor

International Multidisciplinary Conference
"Challenges of Contemporary Higher Education" - CCHE 2024
Kopaonik January 29th - February 3rd 2024
Vol_1

CCHE Challenges of Contemporary Higher Education

Belgrade, 2024

Conference of Academies for Applied Studies in Serbia (CAASS)
University Business Academy in Novi Sad, Faculty of Contemporary Arts, Belgrade

BOOK OF PROCEEDINGS

International Multidisciplinary Conference
"Challenges of Contemporary Higher Education" - CCHE 2024
Kopaonik January 29th - February 3rd 2024
Vol_1

edited by:

Dr Branko Savić, professor

Belgrade, 2024

Impressum
Book of Proceedings, International Multidisciplinary Conference
"Challenges of Contemporary Higher Education" - CCHE 2024
Kopaonik January 29th - February 3rd 2024 https://cche.rs/

Editor
Dr Branko Savić, professor

Publisher
Conference of Academies for Applied Studies in Serbia (CAASS)
Bulevar Mihaila Pupina 2, Novi Beograd, Serbia https://www.caass.rs/
University Business Academy in Novi Sad
Faculty of Contemporary Arts, 12 Svetozara Miletića Street, Belgrade, Serbia
www.fsu.edu.rs

Cover design
Information Technology School ITS -Belgrade

Composition
Conference of Academies for Applied Studies in Serbia (CAASS)
Information Technology School ITS -Belgrade

Print
The Higher Education Technical School of Professional Studies in Novi Sad, Serbia

Circulation
100 copies

Publication year
2024

First edition

ISBN-978-86-82744-02-3 (tom)
ISBN-978-86-82744-00-9 (niz)

WEB AND ANDROID RECOMMENDER SYSTEM FOR STREAMING SERVICES USING APACHE DRILL

Nikola Mekić[138] *Aleksandar Simović* [139]

Abstract: This paper presents the development of a recommender system for streaming services for mobile and Web applications using an interactive data-intensive Apache Drill framework. An innovative approach to utilizing Apache Drill to realize a hybrid recommendation system model by integrating various information technologies is described. A robust and reliable system has been designed and implemented, which, with its functionalities, allows personalized content recommendations of interest to users.

Key words: Apache Drill, PostgreSQL, MySQL, JS, PHP, Android

1. INTRODUCTION

Streaming services have become an indispensable part of the modern digital age. They represent a technology that allows users to access an extensive library of music content without downloading physical copies of music tracks. As with all systems, today's popular services, such as Deezer, Spotify, and iTunes, can improve the personalization and accessibility of the content. The primary goal of this work stems from several critical motives for improving user experience by recommending personalized content of greater interest to users. First, the exponentially fast development of information technologies means that recommendation systems occupy an essential role in different domains and require the integration of different technologies and application methods [1]. Second, low-latency distributed query mechanisms for large data sets, including structured and semi-structured data, require a technological framework applicable to Big Data analytics [2], and crucially, the quality of the personalized recommendation system depends on the integration of multiple data sources, which in many ways can affect both the quality of the personalization service and the user's satisfaction with the received recommendations [3], and the example of Netflix proved to be a critical factor in the "retention" of users on this service [4]. In their paper, Hausenblas and Nadeau describe the functionalities of Drill and its general application [5]. Bløtekjær describes combining the Drill system with other data processing services [6]. At the same time, official documentation presents Apache Drill - A schema-free SQL Query Engine for Hadoop, NoSQL, and Cloud Storage [7].

The primary goal of this work is to create a streaming service that will provide users with a unique and personalized experience by recommending music content of greater interest, both on mobile and on the Web application. Existing music streaming services already tap into personalization to a degree, but Apache Drill opens the door to far deeper individualization. Streaming platforms could tailor recommendations with unprecedented precision by harnessing Apache Drill's capacity for real-time analysis of diverse data sources. This could involve adjusting playlists based on user behavior, environmental context, and even social media trends, all processed in real-time. The result is a listening experience that's not just personalized but truly dynamic, reflecting each user's unique preferences and current situation without resorting to static, one-size-fits-all playlists. Analyzing the application of Apache Drill to create a personalized system of recommendations, we found a minor application of this service. In this paper, we tried to compensate for this deficiency.

The designed system shown in the paper is functional, operational, and available on the Web location [8].

[138] MSc, Information Technology School , Belgrade, Serbia, email: nikolamekic@gmail.com
[139] PhD, Information Technology Professor, Department of Information Technology, Information Technology School, Belgrade, Serbia, email: aleksandar.simovic@its.edu.rs

2. APPLIED TECHNOLOGIES FOR SYSTEM DEVELOPMENT

2.1. Java programming language and Android Studio

Object-oriented high-level programming language Java based on the WORA (Write Once, Run Anywhere) concept, which implies writing Java code on any device, has been implemented to provide the functionality of running software on any device [9]. At the same time, Android Studio is applied as a development environment for creating applications for Android operating systems because it provides support for error recognition, automatic code completion, refactoring, and other functionalities that help in writing programs efficiently [10].

2.2. PHP programming language

PHP, as an open-source server-side scripting language, is used to create Web content because it is focused on working on the server side, where it can manage data, generate dynamic content, and receive and send data [11], which is also its primary function in the realized system. It was chosen for several advantages: it works on most operating systems compatibility with Apache and Microsoft IIS servers, simplicity of syntax, the possibility of working with many databases, and easy integration with HTML [12].

2.3. JavaScript, Ajax, and jQuery

JavaScript, a dynamic object-oriented programming language that forms an indispensable part of Web pages, allowing them to be more than just static documents [13], was used in the realized Web application to design its programming logic. At the same time, AJAX was chosen for functionality such as sending HTTP requests to the server without necessarily refreshing the page [14]. The open-source jQuery library was used for JavaScript functionality and the realization of animations and AJAX requests.

2.4. MySQL and PostgreSQL

One of the main reasons for using MySQL, one of the most famous open-source relational database management systems, in designing a Web application is its compatibility with PHP. It efficiently stores structured data and executes complex SQL queries [15]. In case of a need for greater database flexibility, this system can be replaced or combined with one of the non-relational database management systems such as MongoDB or Cassandra.

Like MySQL, PostgreSQL, as an open-source relational database management system, supports the standard SQL syntax for executing queries but also offers numerous additional functionalities that make it more flexible [15]. Due to the possibility of using transaction logs that enable the recording of all changes in the database and the possibility of data versioning through the functionality of Table inheritance, PostgreSQL was chosen for storing historical data, which is its sole purpose in the designed system.

2.5. Apache Drill

A vital component of the technological framework of the designed system is Apache Drill, a distributed system for interactive ad-hoc analysis of big data. The system can support petabytes of data shared across several thousand servers. The primary goal of Drill is to respond to ad-hoc queries as quickly as possible [5].; supports a wide range of NoSQL database systems, including HBase, MongoDB, MapR-DB, HDFS, MapR-FS, Amazon S3, Azure Blob Storage, Google Cloud Storage, Swift, NAS, and others. Just one query can aggregate data from multiple different sources. For example, it is possible to associate user profiles from a MongoDB database with a directory or data in Hadoop [7].

Apache Drill's prowess in handling vast datasets for personalized experiences finds another excellent application in cloud gaming, where it can optimize streaming quality and tailor game recommendations, ensuring seamless play on any device without the need for high-end hardware [16].

The following chapter provides a detailed description of Apache Drill's operation in the designed and implemented system.

3. DESIGN OF THE PROPOSED MODEL

The amount of data has grown exponentially in the last ten years, as has the number of formats and types [17]. The data is not optimally organized for ad-hoc analyses. Data analysts often encounter various types of data like JSON, XML, and CSV, which are stored in MongoDB, Hadoop, or MySQL, so collecting and organizing them is challenging and time-consuming. Also, when analyzing Big Data, users need to get the result of the requested query in real time without being limited by creating schemas before analyzing the data. With Apache Drill, all of the above is possible. Drill consists of the Drillbit daemon, which can be started online on a local server or one of the nodes within the Hadoop cluster and executes SQL queries on them [6], which makes Drill a completely independent piece of software capable of performing the required task independently of the environment in which it is located.

The key components of Apache Drill are presented on the right in Figure 1:

- **RPC (Remote Procedure Call) endpoint** provides a mechanism for accepting queries from the client and enables communication between the client and Apache Drill. The RPC endpoint ensures that queries are successfully transmitted from clients to the Apache Drill cluster, allowing interaction and sending requests to execute queries.
- **SQL parser** parses the SQL queries sent to Apache Drill. It interprets the query syntax and turns it into an internal representational tree, a logical plan. Usually, the logical plan is present in the memory as Java objects, but it also has a textual form [5]. The SQL parser ensures the correct interpretation of the SQL query and checks that the query is correctly written before it is passed on for processing.
- **Optimizer** is the crucial component that receives the logical plan generated by the SQL parser and optimizes it for the best performance. The optimizer analyzes the logical plan and considers data availability, node distribution, and execution capabilities to generate an efficient physical plan. The physical plan describes the specific operations and scheduling of work on the cluster nodes.
- **Storage engine interface**: Drill is designed to support a variety of data sources, from local hard drives to distributed storage like HDFS. This component provides an interface that allows Drill to access data from various repositories and enables the execution of queries "on the spot" (lat. in-situ) compared to traditional systems. There is no need to load data, create and maintain schemas, or transform data before processing. Instead, it is possible to include a link to a Hadoop directory or a MongoDB data collection in the SQL query [7].

4. SYSTEM OPERATION MODE

In order to demonstrate the effectiveness of the proposed recommender model using Apache Drill, two applications were developed: (I) an application for Android operating systems and (II) a Web application intended for streaming music content. Through both applications, users can listen, rewind, and "like" songs, create playlists, arrange playback orders, and search for songs, artists, albums, and public playlists. Figure 1 shows all components of the designed system:

- **MySQL** – stores data about songs, artists, albums, playlists, and likes.
- **PostgreSQL** – performs storage of data on the history of user activities.
- **Apache Drill** – unifies and displays data from MySQL and PostgreSQL databases.
- **API** – interacts with Android and with the Web application. It allows both applications to enter data into MySQL and PostgreSQL databases and read data from MySQL, PostgreSQL, and Apache Drill.

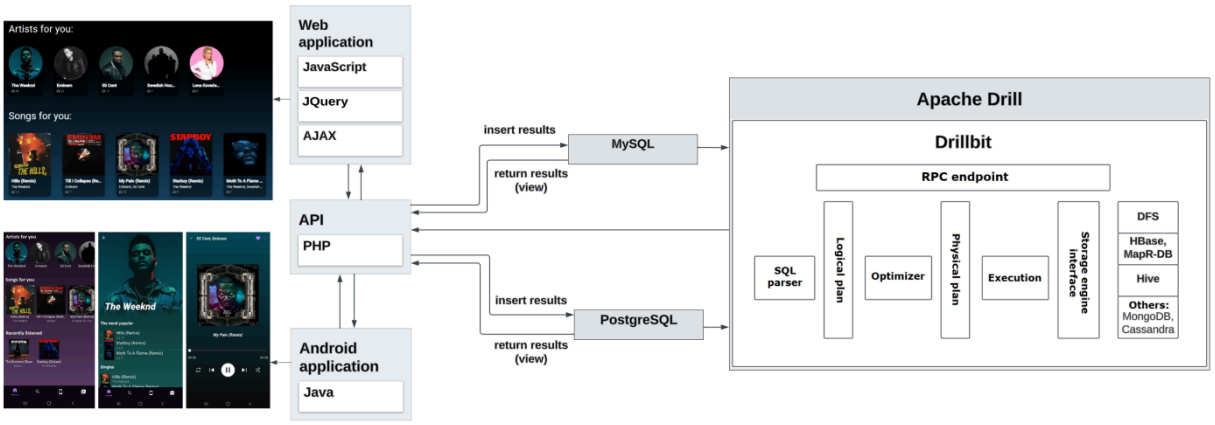

Figure 1 – Block diagram of the realized system [18]

Apache Drill is located locally on an Ubuntu operating system. MySQL and PostgreSQL started, and the JDBC drivers were installed, including the paths where they are located in the configuration section. Figure 2 shows the PHP communication API with Apache Drill that allows song recommendations to the user based on genres and artists. The Client URL PHP library sends a query to Drill via an HTTP request, and if it receives a response, the results are printed via the built-in PHP echo function.

```
select_recommendation_by_artists_and_genre.php:
<?php
session_start();
$user_id =(int) $_SESSION['user_id'];

$ch = curl_init();

curl_setopt($ch, CURLOPT_URL, "http://localhost:8047/query.json");
curl_setopt($ch, CURLOPT_RETURNTRANSFER, 1);
curl_setopt($ch, CURLOPT_POST, 1);
curl_setopt($ch, CURLOPT_HTTPHEADER, array('Content-Type: application/json'));

curl_setopt($ch, CURLOPT_POSTFIELDS, "{\"queryType\":\"SQL\",\"query\":\"APACHE DRILL QUERY\"}");

$result = curl_exec($ch);

if (curl_errno($ch)) {echo "Error";}
else{ echo $result;}
?>
```

Figure 2 – PHP communication with Apache Drill in the function of the recommendation engine

Apache Drill, in the function of the recommendation engine based on the user's listening history from the table in the PostgreSQL database, gives the identifiers (IDs) of the songs the user has not listened to, which are of the same genres and artists. After that, it merges the identifiers with the data from the MySQL database, which contains more detailed information about each song, including song name, artists, genre, path to image display, path to mp3 file, and number of listens.

When the user accesses the home page of Android and Web applications, HTTP requests are sent to the presented PHP code. In the Android application, it was done using the Volley library, while in the Web application, jQuery/AJAX functionality was used. Figure 3 shows sending an AJAX request to a PHP file that prints song recommendations. jQuery accepts JSON data and arranges it into tabs representing songs with their details.

```
function loadRecommendedSongsByArtistsAndGenre(){
  $.ajax({
    url: 'database/select_recommendation_by_artists_and_genre.php',
    dataType: 'json',
    success: function(data) {
      if(data.rows.length>0)
      {
        $.each(data.rows, function(index, song) {

          var cardHtml = '<div class="music-card-recommended-by-artist-and-genre">';

                cardHtml += '<div class="music-image">';
                cardHtml += '<img src="img/' + song.song_img + '" alt="music image">';
                cardHtml += '</div>';

                cardHtml += '<div class="music-details">';
                cardHtml += '<h3 class="music-title">' + song.song_name + '</h3>';
                cardHtml += '<p class="music-artist">'+ song.song_artist +'</p>';
                cardHtml += '</div>';

              cardHtml += '</div>';

          $('.products-list-play-recommended-by-artists-and-genre').append(cardHtml);});
      }
    }
  });
}
```

Figure 3 − Recommendation execution

If the user has listened to at least one song, sections for song and artist recommendations will appear on the home page. Figure 4 shows the user interface of the home pages of the Web and the Android applications. The songs were correctly recommended based on the genres and artists listened to, and only those songs the user had yet to listen to were recommended based on the user's personalized preferences and interests[140].

[140] Program codes are fully covered and, in detail, presented in Mekić's Master's thesis, *Recommender system for streaming services using Big Data and mobile technologies*, Information Technology School in Belgrade, Serbia, 2023 [18].

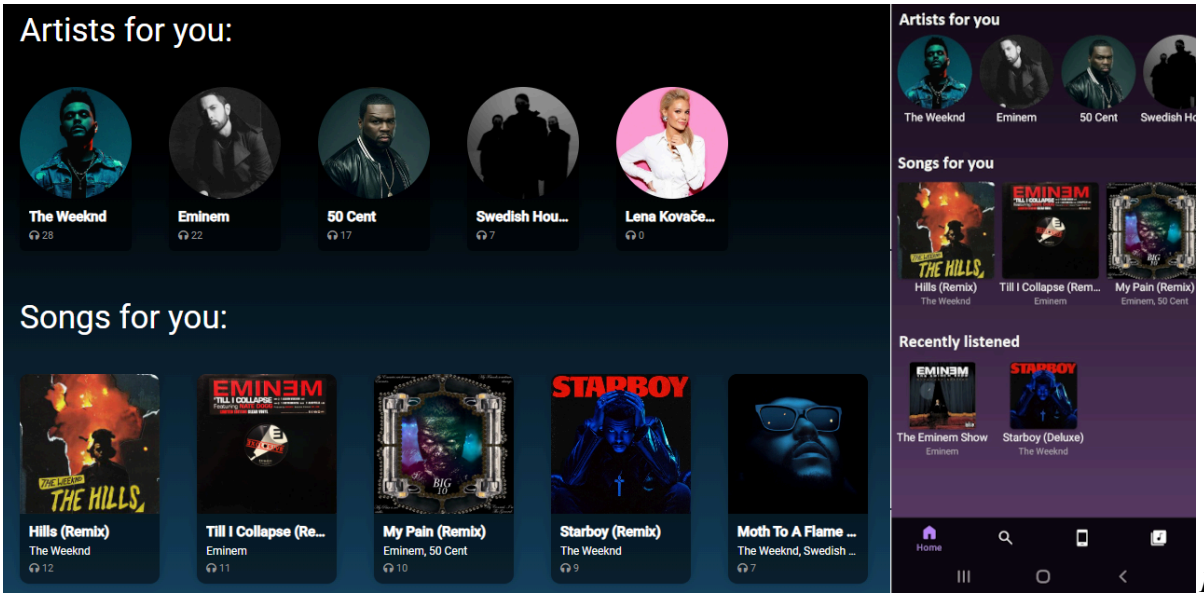

Figure 4 – Home page frontend view for song and artist recommendation – Web application on the left and Android application on the right [8]

5. CONCLUSION

Recommended content shows that using Apache Drill to analyze a large dataset of music content enabled the development of an efficient and scalable recommendation system. What particularly stands out as a contribution of this research is the application of Apache Drill in the context of a recommendation system for streaming services. Previous research has mainly focused on other technologies and tools [1]-[3]. At the same time, this study is, to our knowledge, the first to explore Apache Drill's potential in the recommender systems domain.

The work provides insight into how Drill can effectively process large data sets in the context of personalized music recommendations. Analyzing Apache Drill's performance in this context offers valuable guidance for the music streaming industry and other platforms that want to provide customized music recommendations to their users efficiently. Further research in this area may focus on incorporating additional contextual factors, such as geographic location or mood, to improve the personalization of music content recommendations further.

Exploring the integration of advanced technologies, such as deep learning neural networks, gradient boosting, and collaborative filtering techniques, might enhance the system's efficiency and accuracy in delivering more personalized and context-aware music recommendations.

6. REFERENCES

[1] Simović, A.: Recommender systems in the big data environment using Mahout framework, In *2017 25th Telecommunication Forum (TELFOR)*, Belgrade, Serbia, 2017, pp. 1-4.
[2] Simović, A.: Sistemi preporuke u e-trgovini, *Sinteza 2014-Impact of the Internet on Business Activities in Serbia and Worldwide*, Belgrade, Serbia, 2014, pp. 846-852.
[3] Simović, A.: A Big Data smart library recommender system for an educational institution, *Library Hi Tech*, *36*(3), 2018, pp. 498-523.
[4] Gomez-Uribe, C. A.; Hunt, N.: The netflix recommender system: Algorithms, business value, and innovation, *ACM Transactions on Management Information Systems (TMIS)*, *6*(4), 2015, pp. 1-19.
[5] Hausenblas, M.; Nadeau, J.: Apache drill: interactive ad-hoc analysis at scale, *Big data*, *1*(2), 2013, pp. 100-104.

[6] Bløtekjær, R.: *Performance tuning apache drill on hadoop clusters with evolutionary algorithms*, Master's thesis, 2018.
[7] Apache Drill – Schema-free SQL Query Engine for Hadoop, NoSQL and Cloud Storage, drill.apache.org, Jun 2023.
[8] Music Application for streaming services, nikolamekic.com/muzickaAplikacija, July 2023.
[9] Gosling, J.: *The Java language specification*, Addison-Wesley Professional, 2000.
[10] Android Developers: Android Mobile App Developer Tools, developer.android.com/studio, May 2023.
[11] Welling, L.; Thomson, L.: *PHP and MySQL Web development*, Sams Publishing, 2003.
[12] Reiersol, D.; Baker, M.; Shiflett, C.: *Php In Action, Objects, Design, Agility*, Dreamtech Press, 2007.
[13] Lane, J.; Moscovitz, M.; Lewis, J. R.: *Foundation Website Creation with CSS, XHTML, and JavaScript*, Friends of ED, 2008.
[14] Chaffer, J: *Learning jQuery*, Packt Publishing Ltd, 2013.
[15] Converse, T.; Park, J.; Morgan, C.: *PHP5 and MySQL bible* (Vol. 147), John Wiley & Sons, 2004.
[16] Mekić, N.: Cloud Gaming Services, *Scientific Journal for Contemporary Education and Application of Information Technologies – EdTech Journal*, 2023.
[17] Simović, A.: Big Data Analytics in Public Administration, In *XLIV Symposium on Operational Research SYM-OP-IS 2017*, 2017, Vol. 44, pp. 382-387.
[18] Mekić, N.: *Recommender system for streaming services using Big Data and mobile technologies*, Master's thesis, 2023.

Conclusion

The International Multidisciplinary Conference "Challenges of Contemporary Higher Education" (CCHE) 2024 has successfully convened a diverse group of scholars, researchers, educators, social community members, governmental representatives, employers from the economy sector, and students from around the world. The discussions and research presented have underscored the multifaceted challenges and opportunities that contemporary higher education faces.

Interdisciplinary Collaboration: The conference emphasized the importance of interdisciplinary collaboration in addressing complex challenges. The range of topics, from Medical Sciences to Language for Specific Purposes, showcased that solutions in one field can often be informed and enriched by insights from another.

Global Participation: The hybrid format of the conference ensured that voices from around the world were heard, reinforcing the idea that higher education challenges are global in nature, and solutions must be sought in a collaborative, international context.

Innovation in Higher Education: Numerous papers presented innovative approaches to pedagogy, technology integration, and institutional collaboration. These innovations are pivotal for higher education institutions to remain relevant and effective in a rapidly changing world.

Stakeholder Engagement: A standout feature of the conference was the active participation and collaboration of various stakeholders. Members of social communities provided grassroots perspectives, governmental bodies offered policy-driven insights, employers from the economy sector highlighted practical needs and expectations, and students, the primary beneficiaries of higher education, voiced their aspirations and concerns. This holistic engagement ensures that the strategies and solutions discussed are comprehensive and grounded in real-world needs.

Future Directions: The conference underscored the importance of continuous dialogue and collaboration among all stakeholders. Institutions are encouraged to foster partnerships, both within and outside their disciplines, and with the broader community, to address the evolving challenges of higher education.

Commitment to Excellence: The quality of research and discussions, overseen by the Scientific Committee, showcased a collective commitment to excellence in higher education research and practice.

The CCHE 2024 has set a benchmark for future discussions on higher education. The insights gained and the collaborations formed during this conference, especially with diverse stakeholders, are expected to have a lasting impact on the higher education sector, driving positive change and innovation for years to come.

Warm regards,

Dr Branko Savić, professor

President of the Organizing Committee
International Multidisciplinary Conference “Challenges of Contemporary Higher Education - CCHE 2024”, Kopaonik, Serbia
conference@cche.rs

CIP - Каталогизација у публикацији
Народна библиотека Србије, Београд

004(082)
378(082)
37.091::004(082)

INTERNATIONAL Multidisciplinary Conference "Challenges of Contemporary Higher Education" (2024 ; Kopaonik)
Book of proceedings. Vol 1 / International Multidisciplinary Conference "Challenges of Contemporary Higher Education" - CCHE 2024, Kopaonik January 29th - February 3rd 2024 ; [organisers] Conference of Academies for Applied Studies in Serbia (CAASS), University Business Academy in Novi Sad, Faculty of Contemporary Arts, Belgrade ; edited by Branko Savić. - 1st ed. - Belgrade : Conference of Academies for Applied Studies in Serbia [i. e.] (CAASS) : University Business Academy in Novi Sad, Faculty of Contemporary Arts, 2024 (Novi Sad : Higher Education Technical School of Professional Studies). - 620 str. : ilustr. ; 24 cm

Tiraž 100. - Str. 1-2: Preface / Branko Savić. - Bibliografija uz svaki rad.

ISBN 978-86-82744-02-3
ISBN 978-86-82744-00-9 (niz)

1. Савић, Бранко, 1969- [уредник] [аутор додатног текста]
а) Рачунарска технологија -- Мултидисциплинарни приступ -- Зборници б) Високошколско образовање -- Иновације -- Зборници в) Вештачка интелигенција -- Примена -- Зборници г) Информациона технологија -- Образовање

COBISS.SR-ID 139332105


ISBN: 978-8-68274-402-3