sciendo

ISMSME International Society for Manufacturing, Service and Management Engineering

received: 26 November 2022
accepted: 16 May 2023

pages: 55-71

Trip planning mobile application: a perspective case study of user experience

Ilona Pawełoszek
Jędrzej Wieczorkowski

ABSTRACT
Mobile trip planning applications may contribute to popularising public transport, provided they work efficiently and gain high user acceptance. This article aims to take a closer look at the functioning of the JakDojade application, which has been the most popular platform in Poland for several years, supporting travel planning by public transport. In the presented case study, the authors tried to diagnose problems and indicate the directions of application development. At the same time, through this analysis, the authors aimed to demonstrate the usefulness of researching user comments from the viewpoint of managing the development of mobile applications and related services. A case study methodology was used to perform a descriptive study. Data on user feedback on JakDojade mobile application in Poland comes from Google Play Store. Semantic categorisation of user comments and sentiment analysis allowed for identifying user problems and diagnosing emotions related to its use. The presented methodology allowed for diagnosing typical user problems for the JakDojade application, which may help indicate further development directions. The authors attempted to demonstrate the usefulness of researching user comments from the point of view of managing the development of mobile applications and related services. The semi-automatic approach to text analysis presented in the article highlights the problems related to the study of user reviews. The limitations of the proposed methodology and the possibilities for its improvement were indicated.

KEY WORDS
public transport, mobile application, user experience, UX, sentiment analysis, semantic analysis, software engineering

10.2478/emj-2023-0012

Ilona Pawełoszek

Czestochowa University of Technology, Poland
ORCID 0000-0002-3590-3969
Corresponding author:
e-mail: ilona.paweloszek@pcz.pl

Jędrzej Wieczorkowski

SGH Warsaw School of Economics, Poland
ORCID 0000-0002-1252-8975

Introduction

According to the World Bank Today, 56 % of the world's population lives in cities. This number is expected to grow. By 2050, nearly seven out of ten people will live in cities (World Bank, 2022). The population growth in urban areas will pressure public transport management and infrastructure. The increase in fuel and car prices in the primary and secondary markets causes an increased demand for shared transport models. In the era of universal mobile access to the Internet, mobile applications have become indispensable for younger people as

Pawełoszek, I., & Wieczorkowski, J. (2023). Trip planning mobile application: a perspective case study of user experience. *Engineering Management in Production and Services, 15*(2), 55-71. doi: 10.2478/emj-2023-0012

they support all spheres of personal and professional activity. By offering mobile applications for planning trips, the city can achieve a higher use and economic effectiveness of public transport. Promoting shared transport has many such benefits as (AlKheder, 2021) decreasing the number of private vehicles to ease traffic congestion, resulting in a positive impact on the environment and enhancing the cost-effectiveness, reliability, and accessibility of the road network. The main factors that may discourage passengers from using public transport are:

- Schedule sliding, i.e., the case where vehicles cannot be dispatched according to the planned timetable due to accumulated delays when performing previous trips (Gkiotsalitis & Cats, 2021).
- Problems with purchasing tickets from the electronic ticket machine. In some vehicles, the machines are not present or not functioning correctly. It is not always possible to buy a ticket from the driver, for example, during the restrictions related to the COVID-19 pandemic.
- The difficulty of planning a transfer from one bus or tram to another in the shortest possible time.
- The lack of knowledge about the bus facilities (such as the low floor) essential for people with mobility impairments.
- Uncertainty about the time of arrival and reaching the bus stop.
- Different timetables during the weekends and holidays.

Providing accurate and timely information about public transportation is necessary to increase its dependability. Commuters should know when their trains and buses will arrive so that they can plan their trip accordingly (Rawal, 2018).

The answer to these problems may be a mobile application with real-time access to vehicle information, route plans, and timetables. The authors believe that functional mobile applications meeting user needs are an important factor contributing to the promotion of public transport.

One of the critical success factors of an application is its proper design considering user experience (UX). User expectations are tested during the initial application design process and later, during its production use to apply appropriate corrections and design subsequent versions. User experience research in software engineering usually uses questionnaires, surveys, observations, and psychophysical measurements, often involving users in design work. However, in the operational phase of the application lifecycle, UX research can also be performed by analysing user comments on social networks or in application stores. This, in turn, requires appropriate text mining methods and solving semantic problems, often specific to the analysed language.

This article takes a closer look at the functioning of the JakDojade (How do I arrive?) application, which has been the most popular platform in Poland for several years, supporting travel planning by public transport. By analysing user comments, the authors aimed to diagnose problems and indicate the directions of application development. At the same time, the authors used this analysis to demonstrate the usefulness of researching user comments from the viewpoint of managing the development of mobile applications and related services.

1. Literature review

The debate on the importance of efficient urban transport is not new and has been highlighted in many publications. According to Banister (1995), the attractiveness of particular locations depends on the transport infrastructure's relative accessibility, quality, and quantity. The main problems of efficient transport systems development are (Zavitsas et al., 2010):

- Land use. Well-designed transport networks improve people's accessibility to existing services and amenities. Transport performance is the ratio between the accessibility to certain amenities (including the number of people) by a mode of transport (i.e., how many amenities can be accessed by a specific transport mode in 30 minutes) and the proximity of these amenities (i.e., how many are located in a radius of 8 km) (OECD, 2022).
- Congestion and parking places. Cities are faced with a growing number of cars and limited parking spaces, making people park their vehicles haphazardly on roads and congesting streets.
- Car dependence is among the leading causes of congestion and air pollution. Other consequences are social isolation, discrimination, high expenses, and declining small businesses (Price, 2017).
- Achieving a modal shift from private to public transport, walking, and cycling. In the last few years, the COVID-19 pandemic has caused

a significant impact on public transportation services, travel behaviour, and mode choice preferences (Das et al., 2021).

- Pollution and noise. The EU plans to ban sales of fossil fuel-powered vehicles from 2035 (Abnett, 2022). Many public transport companies in Poland are switching their fleet to electric vehicles. Easily obtained external funds for the purchase of electric buses encourage even the smallest towns to use this solution (Połom & Wiśniewski, 2021).
- Other factors, e.g., political issues and economic prosperity. Urban transport problems can increase travel times and affect the efficiency of commerce, counterbalancing the economic advantages of urban proximity (Zavitsas et al., 2010).

Inner city areas with high building and population densities are problematic in providing efficient transport links and infrastructure while ensuring safety and environmental sustainability. Many cities are thus considering reallocating road space giving more space to public transport and pedestrians (Zavitsas et al., 2010). Due to the spatial and economic problems, local governments have a limited possibility of constant investment in new transport infrastructure. Reducing the share of individual transport will improve the air in the city and the overall quality of life of the population (Bubelíny & Kubina, 2021).

Thus, strategies based on managing existing resources and encouraging the use of public transport seem to be the most appropriate development path. In the information and communication technologies era, solutions and models drawing on the smart city concept have become a hot topic. Great hopes are placed on using various Internet of Things solutions (Shahrour & Xie, 2021). Noteworthy is the selection and integration of IoT tools within smart cities to support decision-making processes in determining investment needs in the public transport lines. It is essential to analyse which IoT tools are available in the city (Stępniak et al., 2021). Implementing real-time locating systems (RTLS) offers the possibility of tracking public transport vehicles.

The basis for creating innovative solutions supporting public services valid for residents is ensuring the openness of source data. Open government data (OGD) is used to develop innovative products, including services and software (Wieczorkowski, 2019). Ensuring the data openness on timetables and information on the current vehicle location enables mobile applications for trip planning.

Mobile applications for paid public services, which usually include public transport, often have to resolve the intermediation problem in the payment-for-service system. The application's popularity may be influenced by implementing appropriate payment solutions with a high level of user trust (Szumski, 2020).

The need to reduce individual transport promotes journey planning platforms involving different public transport means to ease the inhabitant's daily commute. Travel planning applications benefit not only passengers but also the urban transport system and economy. Software companies and startups design and implement these applications, creating new jobs. In addition, Big Data Sources created by these applications may be used for marketing.

However, measuring the benefit of an application for users, i.e., passengers, is problematic. Thus, it is vital to study user experience, particularly the understanding of subjective feelings while using the product (Rota et al., 2009), in this case, a mobile application.

User experience is defined as a person's perceptions and responses that result from the use and/or anticipated use of a product, system, or service (ISO 9241-210, 2010). This formal definition is supplemented by other interpretations: user experience explores how a person feels about using a product, i.e., the experiential, affective, meaningful, and valuable aspects of product use (Vermeeren et al., 2010). Hassenzahl and Tractinsky (2006) distinguished a few prominent perspectives of UX. None of these perspectives fully capture UX, which is about technology that fulfils more than just instrumental needs in a way that acknowledges its use as a subjective. UX is a consequence of a user's internal state, the characteristics of the system (e.g., complexity, purpose, usability, functionality, etc.), and the context within which the interaction occurs.

There are two disparate stances on how UX should be studied, i.e., qualitative versus quantitative, that are not necessarily compatible or can even be antagonistic (Law et al., 2014). UX literature focuses on the previously mentioned methods, such as questionnaires, surveys, and user observation. Such methods often require a manual approach to collecting the necessary data and their subsequent evaluation. As reported (Maia & Furtado, 2016) in a systematic review of UX literature regarding the

automation of such data collection, most studies collected data manually (80 %), 12 % used a mixed form (manual and automated), and 8 % did it automatically. After collecting user experience data, 84 % of the studies performed the user experience evaluation manually, 8 % did it automatically, 4 % used a mixed form, and 4 % did not report the user experience evaluation techniques. Automated UX evaluation is regarded as an unexplored field, especially when adopted for widely used mobile applications (Saleh et al., 2020). However, the authors make such attempts based on user observations. The problem is, therefore, the low level of possible automation of UX research.

The authors of this article focused on the automatic analysis of collected comments. This approach is sporadically mentioned in the literature. Stuart et al. (2015) also examined the reviews of 20 mobile applications from the Google Play Store and the Apple App Store. It was underlined that the unstructured and informal nature of the reviews complicated the analysis. Additional barriers arise from the emotional language enriched with slang, peculiar spelling, and the use of emojis and other symbols (Gimadi, 2021). Many comments address various problems in one review, making it difficult to classify them.

The analysis of user comments (including sentiment analysis) is used when evaluating products, particularly mobile applications. On the other hand, it can also be useful at the city management level to build city development strategies, implement related projects (Jelonek et al., 2020) and evaluate the public transport system (Buran, 2023). The authors of this article focused on the first option.

Semantic analysis is language specific. The abovementioned study concerned comments in English, and the conclusions reached do not need to be valid for other languages. Hence, there is a need for further research on the use of comments for UX assessment, and a research gap here is the use of the Polish language as part of a completely different language group.

The Polish language semantic analysis is a challenging task because of the relatively free word order (Savary & Waszczuk, 2020), which stresses the importance of information rather than following the rules of grammar. Several written corpora of contemporary Polish have been created, which could be used for automatic topic detection. Still, due to copyright issues, they are not freely available for download (Ogrodniczuk et al., 2022).

2. Purpose and characteristics of the JakDojade application

JakDojade application is a public transport planner. Its basic functionality includes searching for the fastest and alternative transport connections and selling tickets online. The application uses city maps, timetables, and user geolocation data. The website is currently available in a mobile version for Android and iOS, as well as in an older Web version.

In 2021, JakDojade was the most popular application for planning trips by public transport in Poland, according to research carried out on behalf of the application developer. Monthly active users (MAU) average about four million. Searching for connections serves over 30 Polish cities or agglomeration areas, and purchasing tickets is possible in about ten cities or regions (Fig. 1). The real-time function is gradually introduced in the application, which consists of using data on vehicles' exact position parallel to the schedule data. So far, it has been implemented in Wrocław, Warsaw, Poznań, Tricity and Kraków. Real-time data improves the quality of finding optimal connections. A passenger can also have a high probability that the selected bus or tram will arrive at the stop at the scheduled time. In addition, the planning considers vehicles' constantly updated speed and travel time on individual sections, not only the timetable values.

However, real-time information requires real-time access to data about the current position of each vehicle. Location data is determined using GPS devices. It is transferred to the transportation company and public transport supervisors and is made available using API programming interfaces. This is still not being done correctly in many cities, resulting in the slow pace of real-time feature deployment. In addition, some vehicles do not report their location accurately due to GPS device failures, problems in data transmission, or driver errors. The quality of the service is also influenced by the frequency of updating the location of vehicles.

The basic version of the mobile application is free, but it provides users with profiled ads. A paid premium version without ads is also available. These advertisements and the sale of user-sourced data are the backbones of the producer's revenue. An increasingly important part of the company's revenues is the intermediation in the sale of online tickets. The

Fig. 1. Geographical coverage of JakDojade application

adopted business model means that the activity is profitable only in larger cities with good advertising and the ticket sales market. Including other centres in the service generates high fixed costs. An alternative is the participation of smaller-town authorities in the costs of including them in the service. The above-described characteristics of the application's operation suggest potential problems affecting user opinions negatively. Therefore, initially, it can be assumed that the criticism and poor opinions may concern:

- speed and reliability of the application operation,
- usability of the application interface,
- the quality of optimisation algorithms,
- news of timetables,
- gaps and errors in real-time information,
- a limited number of cities where the service is provided,
- online ticket sales methods in individual cities,
- problems with paying for tickets,
- the nuisance of displayed advertisements,
- privacy policy,
- prices for the premium version and the profitability of its purchase.

The application evaluation should also be expected to be influenced by the quality of external data downloaded, such as timetables, GPS data of vehicles, and maps. At the same time, the assessment may also be influenced by factors utterly independent of the application, such as the quality of public transport in individual cities and even traffic jams causing delays to timetables. The JakDojade application and business model are described in detail by Wieczorkowski, Chomiak and Pawełoszek (2021). The JakDojade application covers most of the large agglomerations in Poland.

3. Research methods

The great advantage of mobile application stores, such as Google Play, is their star-rating and comments system, allowing users to evaluate the apps and leave feedback. User reviews contain valuable information, such as bug reports, feature requests, and user experiences (Noei & Lynos, 2019). The analysis of user comments is a research method that, in recent years, formed a popular and influential field of

research (Schindler & Domahidi, 2021). This method effectively identifies existing technical problems and other sources of customer dissatisfaction. According to Carter (2022), around 81 % of consumers say they have left a business review around four times a year or less. Therefore, it can be assumed that the analysis of user comments can provide a statistically representative result (Reimer et al., 2021). Although it is increasingly common for customers to leave negative reviews about a product, the most common reason to leave a review is that the service or product was excellent (56 %). The second most common reason to leave a review is that the product was unsatisfactory (41 %) (Carter, 2022). Dissatisfied customers are likely to describe their opinions and experiences in more detail. Sentiment analysis and categorising user comments can indicate the directions for improving the services. Data on user feedback on JakDojade mobile application in Poland was obtained from Google Play Store.

The research procedure consisted of the following steps:

- Gathering comments on the application JakDojade. Comments were extracted from the Google Play store using a Python script and a dedicated library (Mingyou, 2022).
- Preparation of data for analysis in an Excel spreadsheet. The obtained data structure and the description of the variables are presented in Table 1.
- Feature creation, i.e., deriving new features from existing ones. This step involved establishing the gender of commenting users. The gender of users was determined (1) based on their first and last names, using a dictionary method, and (2) based on the words contained in the comment that may indicate gender. In Polish, it is possible to extract past and future verbs and some adjectives indicative of gender. The text analysis was carried out using a VBA script developed by the project authors.
- Calculating descriptive statistics of the data set for individual numerical and categorical variables.
- Application of the CLARIN-PL (Janz et al., 2017) sentiment analysis tool for evaluating the polarity of the comments as positive or negative and also to detect particular emotions accompanying the users of the JakDojade service.
- Separation of thematic categories by using simple text-processing techniques. Here, the first step was to identify the most frequent phrases in the text that could indicate the subject of the comment.
- Tagging the comments and counting the number of comments in each category. Each review could have more than one category.

4. Research results

The first step of the analysis was to identify basic statistics on user ratings and comments. The number of comments and the average assessment from 2011 to the third quarter of 2022 are presented in Fig. 2.

Tab. 1. Variables used in the study

Variable name	Description
Variables in the data set	
reviewId	Unique identifier of the review, string of characters
userName	The user's name and surname or nickname
userImage	User image icon
content	Review text
score	5-point rating scale, where 1 means "Unacceptable" and 5 is "Excellent"
thumbsUpCount	Number of users who found the comment useful
reviewCreatedVersion	Version of app reviewed
at	Date and time of the review
replyContent	Reply text
repliedAt	Date and time of the reply
Derived variables	
gender	User's gender (female, male or unknown)
wordCount	Wordcount in comments
delay	Delay in response to the comment in days
Year	year of adding the comment

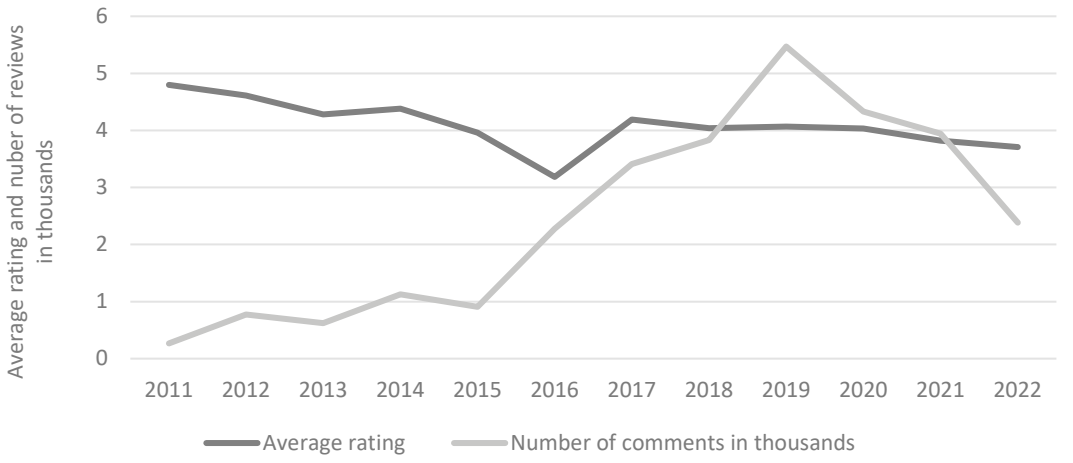

Fig. 2. Number of comments and the average assessment of the JakDojade application

Tab. 2. Descriptive statistics cross-section table

Gender	N	Mean	Std.Dev	Q25	Median	Q75
Female	3584	3.86	1.55	3	5	5
Male	6380	3.66	1.65	2	5	5

Source: elaborated by the author.

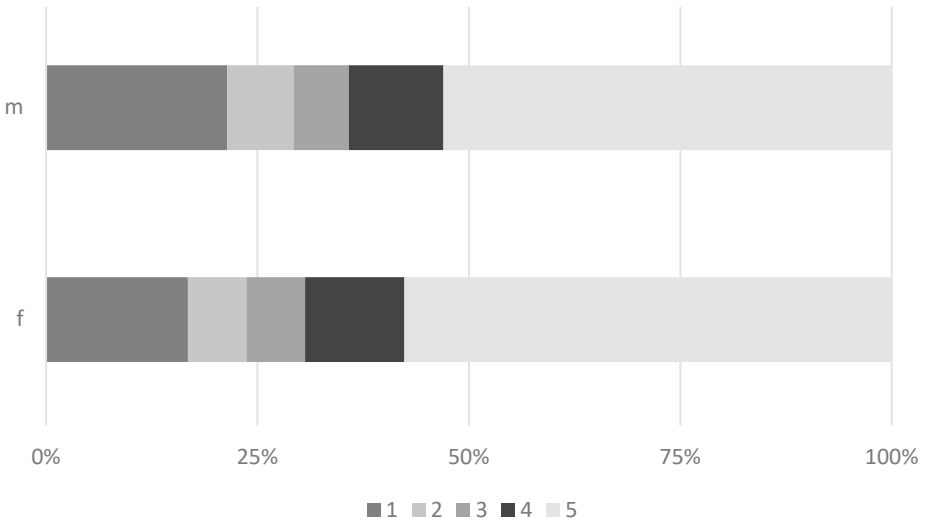

Fig. 3. Share of ratings 1–5 by gender

The average user ratings over the years ranged from 3.18 to 4.8. Notably, the median in each year of the analysis was five, meaning that over half of the users rated the application very well. The lowest average rating was recorded in 2016 (3.18). Then, the ratings of the JakDojade application broken down by gender were analysed. By analysing the names and surnames of commentators and the grammatical analysis of statements, the authors were able to determine the gender of 33 % of commentators.

The authors decided to check whether men and women differed in the assessment of the application. The obtained results are presented in Table 2. The chart in Fig. 3 presents the visualisation of the number of grades on a scale of 1–5, broken down by gender.

The statistics analysis on the length of comments (the variable wordCount) showed that women write longer reviews. However, the difference is not significant and likely has arisen by chance. For men, the

Tab. 3. Descriptive statistics on comment length

Gender	N	Mean	Std.Dev	Q25	Median	Q75
Female	3584	16.14	17.73	3	10	22
Male	6380	15.82	18.75	3	9	21

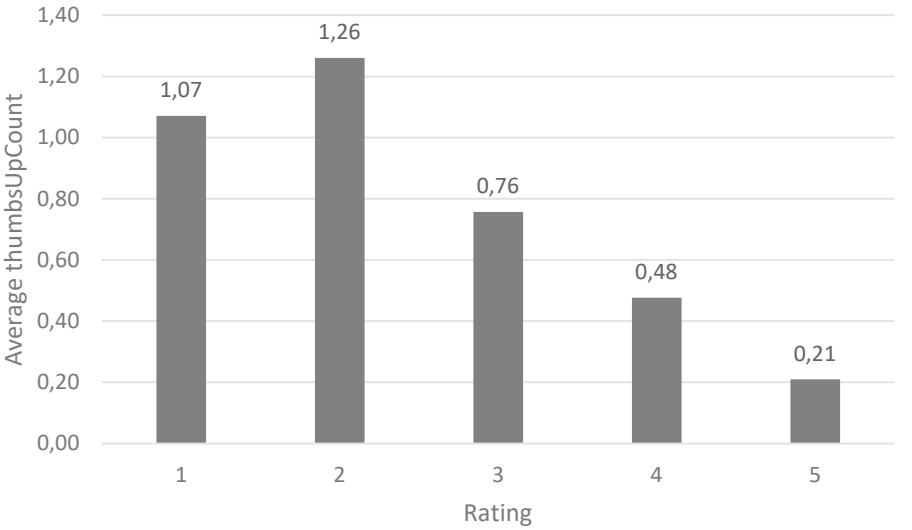

Fig. 4. Average variable thumbsUpCount by application rating

median and average are 15.81 and nine words, and for women, it is 16.14 and ten, respectively. Detailed statistics on comment length by gender are presented in Table 3.

The Google Store comment system allows users to evaluate the usefulness of a review. Users can click the thumb icon if they find a comment valuable. In the analysed dataset, the usefulness of the comment is measured by the variable thumbsUpCount. To take a closer look at the usefulness of the comments, the authors examined the following relationships:

- The usefulness of comments, depending on their length.
- In this study, the authors assumed that the most valuable comments were those with the most text and detailing why users were satisfied or dissatisfied. However, this assumption was rejected due to the correlation coefficient between wordcount and thumbsUpCount amounting to 0.19.
- The usefulness of comments, depending on their rating.

The correlation coefficient between these two variables was -0.01, showing no linear relationship. Fig. 4 presents the average number of thumbs by rating.

On average, the comments that rated the app low (one and two stars) proved to be the most helpful.

The analysis of the length of comments broken down by star rating showed that these comments contained the most text (Fig. 5).

Considering the usefulness and length, comments evaluating the application with one, two, and three stars were selected for further analysis. These comments detailed the reasons for user dissatisfaction and the application use barrier.

Responding to comments is crucial to building a customer-oriented marketing strategy. By timely response, the company will improve customer retention rates and loyalty. Answering the comments also increases the chances of acquiring new customers by showing the company's commitment. The authors decided to check the response time and the share of answered comments according to ratings. In the research sample, the company responded to 64 % of the comments. The share of replied comments broken down by rating is presented in the chart in Fig. 6. Detailed statistics on the reply time are presented in Table 4.

In most cases outside the 4-star rating, the median response time to comments was one day. On average, the company responds to 1-star comments the fastest. When examining the dependence of the response delay on the comment length, the authors used the correlation coefficient. Its value was -0.00128,

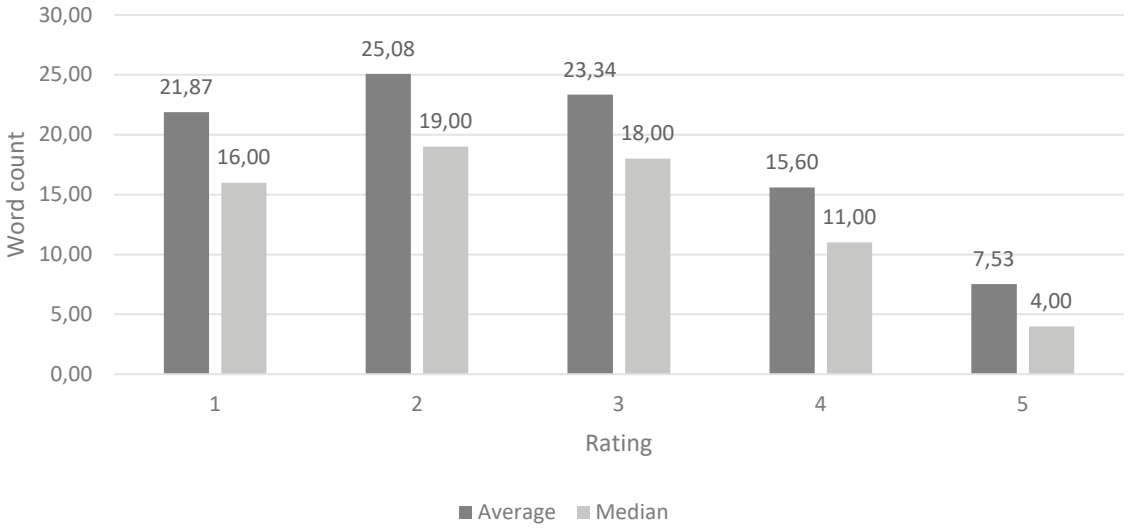

Fig. 5. Average and median word count by application rating

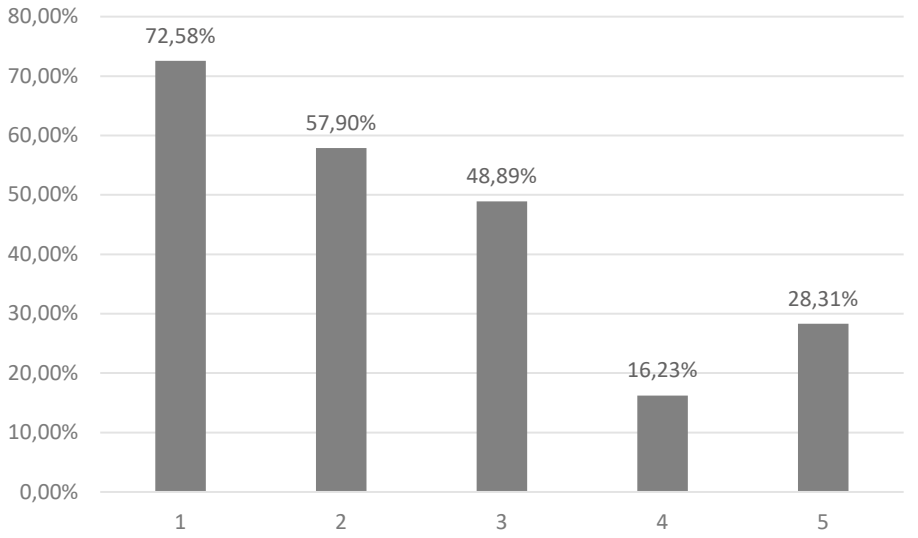

Fig. 6. Percentage of replied comments by rating

Tab. 4. Descriptive statistics of application ratings

Rating	N	Average	Median	Min	Max	Q1	Q3	Std. dev.
1	3041	4.68	1.00	0.00	122.00	0.00	5.00	9.17
2	1088	5.07	1.00	0.00	116.00	0.00	5.00	10.46
3	1088	5.07	1.00	0.00	116.00	0.00	5.00	10.46
4	614	8.57	2.00	0.00	182.00	0.00	8.00	18.07
5	4970	5.49	1.00	0.00	165.00	0.00	4.00	16.16

which indicates no relationship between the two variables.

To diagnose the problems and barriers faced by users of the JakDojade mobile application, sentiment analysis and semantic analysis were used to distinguish the categories of emerging issues. The sentiment analysis using the CLARIN-PL tool was applied to identify the users' emotions. Due to a large number of comments (29.343) and limited computing power, the sentiment analysis was performed on a sample of 400 comments. Therefore, the sample contained 80 reviews from each rating 1–5, respectively. The graph

presented in Fig. 7 clearly shows the advantage of positive emotions.

A more detailed analysis with the CLARIN-PL tool allows for extracting phrases related to various emotions, both positive (such as joy, trust, and looking forward to something unexpected) and negative (sadness, anger, fear, disgust, and negative surprise). The results of the analysis are presented in Fig. 8.

Among the identified positive emotions, joy and trust prevailed. The content analysis of positive comments showed that users were happy to plan their trip well with the help of the application. They trusted the application, as evidenced by numerous phrases proving no experienced disappointment. Sadness and anger predominated among negative emotions. These emotions were related to many shortcomings of the application's functioning. To find out the causes, a semantic analysis of the text was made and is described in the further part of the article.

A more in-depth analysis with the CLARIN tool extracts the emotional tinge of the application characteristics mentioned by users in the comments. The summary of the analysis results is shown in Fig. 9.

The results are largely in line with the emotions identified earlier. The most frequently indicated feature is the usability of the application. Users often described cases of successfully reaching destinations or not getting lost in an unfamiliar city. Among the negative traits, the most frequently mentioned were uselessness, harm and misfortune.

It is possible to identify problems by analysing comments with low ratings, i.e., one, two or three stars. The first step was identifying recurrent topics. For this purpose, the 100 most extended negative

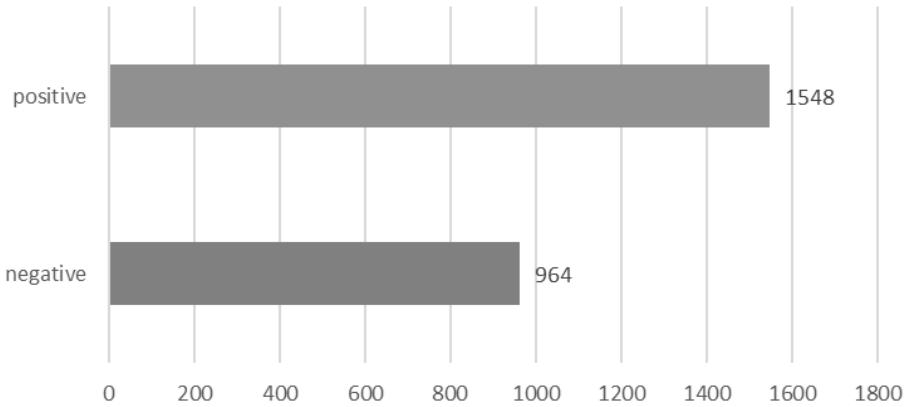

Fig. 7. Polarity of comments

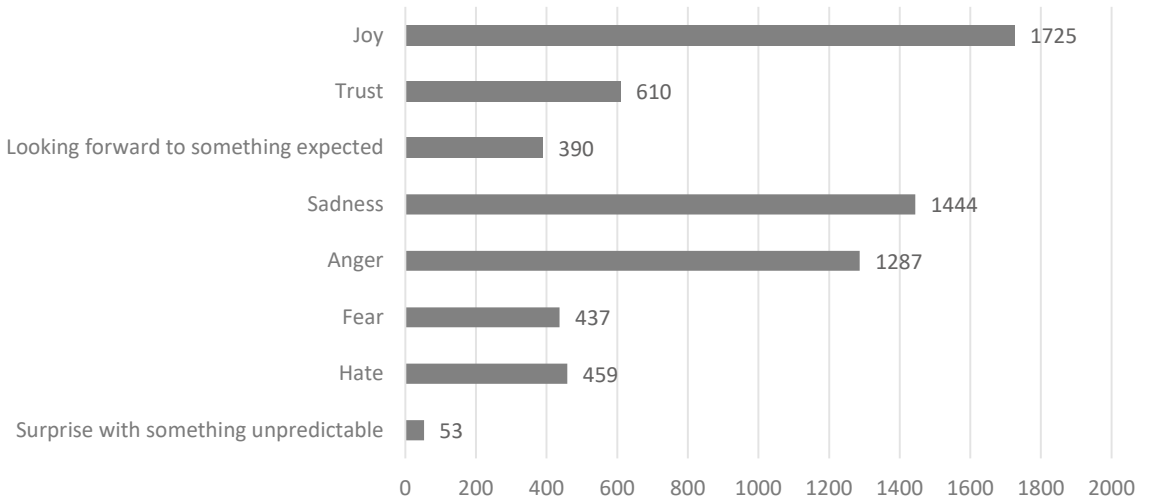

Fig. 8. User emotions detected by sentiment analysis

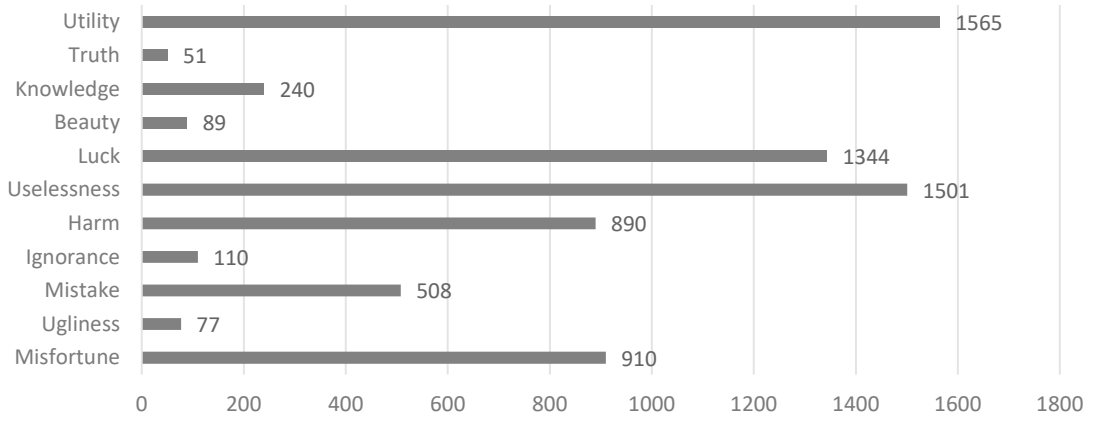

Fig. 9. Emotional colouring of the features of the JakDojade mobile application detected by the sentiment analysis

Tab. 5. Categories and key phrases used for tagging user comments

Category	Description	Key phrases
TimetableS	Problems with the timeliness of the timetable, updating, and searching for optimal connections	Timetable, showing delay, showing connections, change of means of transport, hours, searching, delay, route, train, timetable update
Sharing personal data	The application collects the personal data of users, for which they must consent to use the application	Statute, policy, personal data, agreement, private, protection, consent
Energy consumption	Energy consumption by the application, cases of power drain	Battery, energy consumption, power drain
Finance	Payments for tickets, payment methods, complaints about transactions, payments for the premium application	Paid, payment, return, money, scam, wallet, complaint, ticket
Navigation	Problems with in-app navigation, road showing on map and GPS location	Navigate, navigation, GPS, location, map update
Interface	Application interface, ease of use, interface changes, suggestions for changes	Appearance, interface, font, button, readable
Premium app	Comments related to the app's premium version and displaying advertisements in the non-premium version. Comparison of free and premium applications	Ads, premium
Application	Problems with using the mobile application and updating.	Efficient, stable, suspending, mule, silt, slime, jam, update
Cities	Reporting problems related to specific places, service in cities	City, town, names of the cities
Holiday	Problems encountered with revised timetables	Holiday, weekend, Saturday, Sunday

comments sample was selected and analysed to find the most frequent issues reported by users. Next, the authors extracted the concepts that indicated problems and their synonyms; this way, 54 keywords were identified and categorised. Table 5 shows the categories of topics. Due to the declension rules in the Polish language (inflexion of a word by cases and numbers), a simple searching method did not give accurate results. To avoid this problem, the morphemes of the words (a part of the word with a recognisable meaning) were used. The corpus of all the comments was searched to find the morphemes and count their occurrences in the text.

In the case of key phrases consisting of two words, an original VBA function was used to examine the close coexistence of morphemes in the text. Topic extraction from the text was made by automatically tagging the reviews according to predefined categories.

The frequency of words in particular categories is presented in the diagram in Fig. 10. The overwhelming number of negative comments criticise the time-

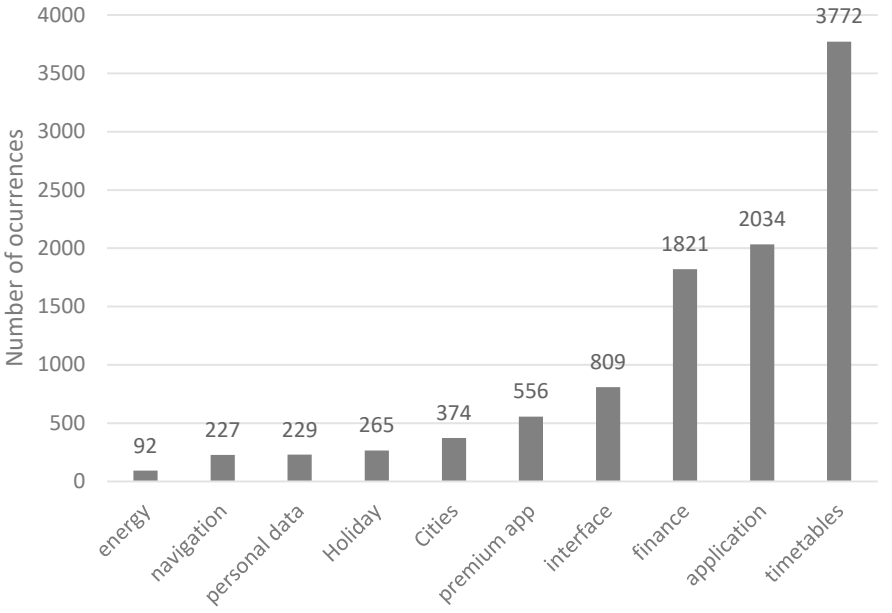

Fig. 10. Number of problem occurrences in user comments by category

	Energy consumption	Holiday	Sharing personal data	Cities	Premium version	Navigation	Interface	Finance	Application	Timetables
Energy consumption		0,00%	0,03%	0,01%	0,04%	0,08%	0,08%	0,03%	0,19%	0,08%
Holiday			0,01%	0,04%	0,00%	0,04%	0,03%	0,10%	0,20%	1,27%
Sharing personal data				0,06%	0,15%	0,34%	0,19%	0,45%	0,45%	0,94%
Cities					0,11%	0,44%	0,29%	0,50%	0,46%	1,22%
Premium version						0,76%	0,75%	0,61%	0,93%	1,37%
Navigation							1,39%	0,39%	1,74%	2,48%
Interface								0,54%	2,23%	2,92%
Finance									2,88%	2,83%
Application										7,59%
Timetables										

Fig. 11. Co-occurrence of the topics matrix — support coefficients

table. This is a particularly important problem when planning a journey with transfers. In large cities, the application supports various means of transport, such as buses, trams, metro and light rail. Updating these timetables may vary by carrier.

Another frequently mentioned problem is the functioning of the application. In particular, software stability and crashes, which, in turn, are related to the inability to make payments. This situation may create a legal problem and expose the passenger to travel without a purchased ticket.

The coexistence of topics was determined using the association rules method known from the market basket analysis in the marketing field. It analyses co-occurrence patterns and determines the strength of the link between two items in the dataset (Szymkowiak et al., 2018).

Every comment rated from one to three was tagged with topic categories. The analysis consisted of identifying the frequency of coexisting categories in user reviews. This approach allowed for extracting support and confidence coefficients for the recognised rules. Support emphasises how popular an item set is, and confidence denotes the likelihood of certain items occurring together. The matrix in Fig. 11 shows support coefficients for the identified two-item rules.

If x and y are the categories listed in the first row and column of the table, respectively, the support

y/x	Energy consumption	Holiday	Sharing personal data	Cities	Premium version	Navigation	Interface	Finance	Application	Timetables
Energy consumption		0,00%	2,17%	1,09%	3,26%	6,52%	6,52%	2,17%	16,30%	6,52%
Holiday	0,00%		0,38%	1,13%	0,00%	1,13%	0,75%	3,02%	6,04%	38,11%
Sharing personal data	0,87%	0,44%		2,18%	5,24%	11,79%	6,55%	15,72%	15,72%	32,75%
Cities	0,27%	0,80%	1,34%		2,41%	9,36%	6,15%	10,70%	9,89%	25,94%
Premium version	0,54%	0,00%	2,16%	1,62%		10,97%	10,79%	8,81%	13,31%	19,60%
Navigation	2,64%	1,32%	11,89%	15,42%	26,87%		48,90%	13,66%	61,23%	87,22%
Interface	0,74%	0,25%	1,85%	2,84%	7,42%	13,72%		5,32%	22,00%	28,80%
Finance	0,11%	0,44%	1,98%	2,20%	2,69%	1,70%	2,36%		12,63%	12,41%
Application	0,74%	0,79%	1,77%	1,82%	3,64%	6,83%	8,75%	11,31%		29,74%
Timetables	0,16%	2,68%	1,99%	2,57%	2,89%	5,25%	6,18%	5,99%	16,04%	

Fig. 12. Confidence matrix of co-occurring topics
Source: elaborated by the author.

factor S x, y located in the cell at the intersection of the row and column can be written as follows:

S(x,y) = number of comments containing both x and y / total number of comments

The confidence coefficients (Fig. 12) at the top right of the matrix were calculated using the following formula:

Confidence (X→Y) = (Support (X,Y)) / (Support (X))

Accordingly, the confidence coefficients at the left bottom part used the formula:

Confidence (Y→X) = (Support (Y,X)) / (Support (Y))

The higher the confidence, the greater the likelihood that the comment that includes topic X will also include topic Y.

Knowledge of support and confidence in the rules may indicate a semantic connection between the considered problem categories (Zhuge et al., 2004). The three and more-component rules were also distinguished; however, their supports were very low, less than 1 %; therefore, they were considered insignificant and omitted in the inferential analysis.

5. Discussion of the results

The presented analysis shows that the number of comments grew dynamically from 2015 to 2019 when the application began to gain recognition from increasingly more users. This fact can be explained by expanding the action zone with new cities and adding new functions. At the beginning of 2015, 20 new cities, suburban zones and railway lines were added to the application. However, major cities were already served much earlier. Gradually, in the year 2008, Poznań was introduced to the service. In the following years, the area of operation was extended to the next seven large urban areas: Wrocław, Warsaw, Szczecin, Kraków, Łódź, Bydgoszcz and Toruń. For about ten years, the offer was extended to significantly smaller cities. Thus, the number of comments was more related to the popularisation of the service and not its availability in other cities.

The decrease in the number of users in 2020 was caused by restrictions related to the COVID-19 pandemic. During the pandemic, public transport was considered a contamination hazard, and the number of passengers decreased significantly. Additionally, timetables were limited to the necessary minimum. The scope of the analysed data does not allow conclusions about the application's popularity during the period of easing the restrictions after the pandemic.

In addition to the development of the application by adding new cities, the range of offered services increased (including the possibility to purchase tickets and track the location of vehicles in real-time). Their evaluation may be of particular interest to the application provider. The clear increase in the number of comments in 2015–2019 can be associated with extending the application's functionality.

The average app ratings in the analysed period slightly decreased. However, as already noted, the median in each analysed year was five, and more than half of the users assessed the app very well. The lowest average rating value recorded in 2016 was likely related to the increasing complexity of the application and growing user expectations. It should be noted that other platforms with similar functions were also developing in parallel, the most competitive of which

was Google Maps. In 2015, Google Maps introduced the possibility of using offline maps and many new functions.

It is more difficult to interpret the differentiation in the number and value of rates by gender. Especially for more than half of the comments, it was impossible to assign a gender, so there are no grounds to conclude that one gender does not dominate in this group. Undoubtedly, however, in the case of comments with the identified gender of the commentator, women rated the application higher than men by giving excellent ratings (5) and fewer poor ratings (1). The authors observed a similar relationship in another study, i.e., the analysis of comments from car-sharing applications (Pawełoszek, 2022). Therefore, it can be suspected that this is a general rule resulting from the psychological characteristics of gender, although it would require further research. Similarly, it is difficult to explain some differences in the length of written comments.

The analysis of the usefulness (thumbs up) of comments depending on the number of stars given to the comment indicated that the most useful comments were negative (one or two stars), indirect neutral comments (three stars), and the least useful were positive comments (four or five stars). This is understandable as the content analysis of the comments showed that these worse comments were more likely to indicate specific bugs and deficiencies of the application. Positive comments, in turn, generally praised the application only by confirming the correct operation, i.e., by confirming what should be the norm. Consequently, a detailed analysis of negative comments may be of significant value to the application developer. The creator of the JakDojade application responds to negative comments most often and quickly. In addition, other users can see the answer from the point of view of the application's developer. Such a strategy also aims at building the image of a company that cares about good contact with users and their loyalty.

From the company's viewpoint, the comment sentiment analysis at the most general level of identification of positive and negative emotions may have a complementary role to the analysis of the system's comments rating based on the number of stars. The sentiment analysis, like the number of stars, allows for observing trends in product evaluation. In the case of the tested application, the predominance of positive over negative emotions confirmed the observed high mean and median rating in the star system. Further, it is possible to isolate positive and negative emotions. In practice, however, it is difficult to identify the cause of a given feeling without analysing the content of the entire specific comment.

More practical information can be obtained from the semantic features' analysis of the tested application. However, it should be considered that for mobile applications, analogous but opposite features will appear both positive and negative, as in the case of "utility" and "uselessness". In this case, usability was assessed highly differently. In the case of the JakDojade application, the frequent occurrence of the "mistake" feature was also disturbing, which is probably related to the inconsistency of the provided timetables and proposed routes with the actual functioning of transport. It should be noted that sentiment analysis can be carried out (as in his study) to a large extent with the help of ready-made tools. On the one hand, this significantly simplifies the analysis, but on the other hand, the tool is then universal and not adapted to solving a specific problem. Consequently, the usefulness of such an analysis is often limited.

The procedure is different in the case of identifying typical problems appearing in the comments. The used approach required expert identification of the anticipated problem categories and assigning them to keywords. Such a procedure is often iterative, based on repeated attempts in the event of imminent failure. Keyword selection with this approach is partially subjective. The approach requires knowledge of the tested product or service and is associated with a greater amount of work, but the analysis is tailored to the solution of a specific problem. Ultimately, the authors distinguished ten categories for the tested application, assigning them several keywords. Some of the categories were universal for the evaluation of various mobile applications. Such an important category of problems typical of all mobile applications are those related to the stability, performance, and updates of the application (the "application" category). In the example used by the authors, as expected, many negative comments contained keywords assigned to this category. There were often other comments related to the application's features belonging to another distinguished category of "interface" and much less frequently to the category of "energy consumption".

The categories typical of most mobile applications also include "sharing personal data". In the discussed case, the keywords related to this category did not often appear in negative comments. Nevertheless, the business model of many mobile application providers, including JakDojade, is based on personal

data processing, which may raise user objection, and the level of such objection should be monitored.

A similar situation occurs when offering different versions of an application, usually the basic free version and the advanced paid premium version. This is also the case with JakDojade. The negative comments included several keywords assigned to the "premium version" category, which may indicate the dissatisfaction of users with such a pricing policy or a malfunction of the paid version. On the other hand, such comments may indicate the limitations and shortcomings of the basic product version, including the nuisance of displayed ads.

Another group of categories is related to the specificity of the application or related services. In this case, such a category is primarily "timetables". In negative studied comments, it was by far the most important category, which is in line with previous assumptions. The comments relate to incorrect and outdated timetables, inconvenient proposed routes, and incomplete real-time information. This category is related to another distinguished category concerning holiday schedules (the "holiday" category). Another important category was related to the application's basic functionality, i.e., the purchase of tickets (the "finance" category). The comments analysis shows that errors resulting from unsuccessful ticket purchase transactions using the application are extremely bothersome.

A characteristic feature of the JakDojade application is its diversified usability in individual cities due to the specificity of local public transport. Hence, the comments had keywords assigned to the "cities" category. This is mainly due to the location of the existing problems in the geographic comments (usually related to timetables or the purchase of tickets). However, contrary to the authors' expectations, the number of such keywords was not significant. The application usually does not receive comments from people from cities where the service has not been implemented. Hence, there are few entries for the limited number of cities where the service is provided.

Some keywords may be difficult to interpret due to their ambiguity. For example, "navigation" can, on the one hand, be understood in the context of geographical navigation related to the user's location and the location of individual public transport vehicles. On the other hand, it can also be understood in the context of navigation within the user interface. Here, the analysis of the coexistence of keyword pairs or their categories, as well as the confidence analysis, can be useful. The co-occurrence results (Fig. 11) showed a strong relationship between the "navigation" category with both the "interface" category and the "timetables" category. At the same time, the confidence analysis (Fig. 12) showed that a significant part of the comments relating to the category "navigation" referred simultaneously to at least one of the categories: "interface", "application", or "timetable". Such links help to understand the context of the keywords' occurrence.

However, some of the strong co-occurrences between the most popular categories (e.g., the most popular "application" and "timetable" categories) are trivial, generally related to the intended use of the application, and, in practice, not interpretable. It is also possible to analyse the coexistence of three or more categories, i.e., either those with the strongest support or expertly selected combinations of several categories. This is justified when it is difficult to interpret the presence of individual categories or their pairs.

The above analyses can at least partially be automated thanks to the semantic analysis of the text. Unfortunately, in practice, it is necessary to examine the content of individual comments in parallel to understand their meaning more precisely. The authors had almost 8 000 comments at their disposal. The random selection showed that they often described specific situations in which the application did not work properly (e.g., timetables, purchase of tickets, updates of the application or timetables, operation of the application on specific devices, the complaint handling process, etc.). It is an important source of knowledge for application developers, but this type of comment analysis is very laborious. In addition, many comments contain numerous typos, incorrect abbreviations, and other linguistic errors, making it a challenge for automated analysis.

Conclusions

The research confirmed the usefulness of analysing comments on mobile applications posted in application stores. Such analysis helps application developers identify and understand user problems. However, the possibilities of automating this type of analysis are very limited. Simple tools that do not require a significant amount of work, such as the analysis of the ratings given to the application by its users or automated sentiment analysis, do not sufficiently explain the motivations of the evaluators. The semantic analysis of the content of comments gives

much greater possibilities. Still, it is associated with a more significant workload resulting from the need to select the appropriate keywords and their proper categorisation, which is related to the amount of work but also requires a good knowledge of the tested application at the beginning. In practice, if an application developer expects detailed information about user feedback, they usually have to refer to the content of specific opinions. Their analysis can, to some extent, be automated thanks to text mining tools, but it already requires much work.

The authors' initial intent was to question the extent to which the analysed comments would concern the application and the entire service. In the case of JakDojade, comments could relate to the quality of public transport in general, timetables, punctuality, and other factors independent of the application. However, the analysis of the comments showed that the vast majority of entries concerned the application and not its broader context. This is an important conclusion from the viewpoint of the usefulness of the proposed method for managing the software development and maintenance process.

Therefore, it was possible to diagnose typical user problems for the selected JakDojade application, which may help indicate further development directions. However, most of all, it was possible to demonstrate some, though partially limited, usefulness of the proposed methods of researching user comments from the point of view of managing the development of mobile applications. It should be assumed that similar methods can also be used in the case of other products and services, although it should be noted that access to comments related to mobile applications is extremely convenient due to their distribution through online stores with extensive comment systems.

Literature

Abnett, K. (2022). EU lawmakers back ban on new fossil-fuel cars from 2035, *Reuters*. Retrieved from https://www.reuters.com/business/autos-transportation/eu-lawmakers-support-effective-ban-new-fossil-fuel-cars-2035-2022-06-08/

AlKheder, S. (2021). Promoting Public Transport as a Strategy to Reduce GHG Emissions from Private Vehicles in Kuwait. *Environmental Challenges*, *3*, 100075. doi: 10.1016/j.envc.2021.100075

Banister D. (1995). *Transport and Urban Development*, London, UK: E & FN Spon.

Bubelíny, O., & Kubina, M. (2021). Impact of the concept smart city on public transport. *Transportation Research Procedia*, *55*, 1361-1367. doi: 10.1016/j.trpro.2021.07.120

Buran, B. (2023). Sentiment Analysis on Public Transportation During Covid-19: An Exploratory Study. In F. Calisir (Ed.), *Industrial Engineering in the Age of Business Intelligence. GJCIE 2021. Lecture Notes in Management and Industrial Engineering* (pp. 119–132). Cham: Springer. doi: 10.1007/978-3-031-08782-0_10

Carter, R. (2022). *The Ultimate List of Online Review Statistics for 2022*, B2B Insights. Retrieved from https://findstack.com/online-review-statistics/

Das, S., Boruah, A., Banerjee, A., Raoniar, R., Nama, S., & Maurya, A. K. (2021). Impact of COVID-19: A radical modal shift from public to private transport mode. *Transport Policy*, *109*, 1-13. doi: 10.1016/j.tranpol.2021.05.005

Gimadi, D. (2021). Web-sentiment analysis of public comments (public reviews) for languages with limited resources such as the Kazakh language. In *Proceedings of the Student Research Workshop Associated with RANLP 2021* (pp. 65–68). INCOMA Ltd.

Gkiotsalitis, K., & Cats, O. (2021). At-stop control measures in public transport: Literature review and research agenda. *Transportation Research Part E: Logistics and Transportation Review*, *145*, 102176. doi:10.1016/j.tre.2020.102176

Hassenzahl, M., & Tractinsky, N. (2006). User experience - a research agenda. *Behaviour & Information Technology, 25*(2), 91-97. doi: 10.1080/01449290500330331

ISO 9241-210:2010. Ergonomics of human system interaction -Part 210: Human-centred design for interactive systems.

Janz, A., Kocoń, J., Piasecki, M., & Zaśko-Zielińska, M. (2017). *plWordNet as a Basis for Large Emotive Lexicons of Polish*. LTC'17 8th Language & Technology Conference. Poznań, Poland. Retrieved from https://www.researchgate.net/publication/322684200_plWordNet_as_a_Basis_for_Large_Emotive_Lexicons_of_Polish

Jelonek, D., Stępniak, C., Turek, T. & Ziora, L. (2020). Planning Cities Development Directions with the Application of Sentiment Analysis. *Prague Economic Papers*, *3*, 274-290. doi: 10.18267/j.pep.740

Law, E., van Schaik P., & Roto V. (2014). Attitudes towards user experience (UX) measurement. *International Journal of Human-Computer Studies*, *72*(6), 526-541. doi: 10.1016/j.ijhcs.2013.09.006

Maia, C.L.B., & Furtado, E.S. (2016). A Systematic Review About User Experience Evaluation. In A. Marcus (Ed.) *Design, User Experience, and Usability: Design Thinking and Methods*. DUXU 2016. *Lecture Notes in Computer Science*, 9746. Cham: Springer. doi: 10.1007/978-3-319-40409-7_42

McIlroy, S., Ali, N., Khalid, H., & Hassan A. E. (2016). Analyzing and automatically labeling the types of user issues that are raised in mobile app reviews. *Empirical Software Engineering*, *21*, 1067-1106. doi: 10.1007/s10664-015-9375-7

Mingyou, J. (2022). *Google-Play-Scraper*. Retrieved from https://pypi.org/project/google-play-scraper/

Noei, E., & Lyons, K. A. (2019). A survey of utilizing user-reviews posted on Google play store. In *Proceedings*

of the 29th annual international conference on computer science and software engineering 2019, (pp. 54–63). IBM Corp., USA, CASCON '19.

OECD (2022). *Efficient land use and public transport systems for sustainable cities (SDG 11)*. Retrieved from https://www.oecd-ilibrary.org/sites/79c1e03d-en/index.html?itemId=/content/component/79c1e03d-en

Ogrodniczuk, M., Pęzik, P., Łaziński, M., & Miłkowski M. (2022). *European Language Equality. D1.27 Report on the Polish Language*. Retrieved from https://european-language-equality.eu/wp-content/uploads/2022/03/ELE___Deliverable_D1_27__Language_Report_Polish_-1.pdf

Pawełoszek, I. (2022). Towards a Smart City—The Study of Car-Sharing Services in Poland. *Energies*, *15*, 8459. doi: 10.3390/en15228459

Połom, M., & Wiśniewski, P. (2021). Implementing Electromobility in Public Transport in Poland in 1990–2020. A Review of Experiences and Evaluation of the Current Development Directions. *Sustainability*, *7*, 4009. doi:10.3390/su13074009

Price, A. (2017). *The Negative Consequences of Car Dependency, Strong Town*s. Retrieved from https://www.strongtowns.org/journal/2017/11/2/the-negative-consequences-of-car-dependency

Rawal, P. (2018). *Public transport – A key element of smart cities*, Observator Research Foundation. Retrieved from https://www.orfonline.org/expert-speak/43591-public-transport-a-key-element-of-smart-cities/

Reimer, J., Häring M., Loosen W., Maalej W., & Merten L. (2021). Content Analyses of User Comments in Journalism: A Systematic Literature Review Spanning Communication Studies and Computer Science. *Digital Journalism*. doi: 10.1080/21670811.2021.1882868

Roto, V., Obrist, M., & Vaananen-Vainio-Mattila, K. (2009). *User experience evaluation methods in academic and industrial contexts*. In *Proceedings of the Workshop UXEM* 9 (pp. 1–5). International Conference on Human-Computer Interaction.

Saleh, A. M., Al-Badareen, A. B., & Enaizan, O. (2020). Automated user experience tool development for mobile application. *International Journal of Advanced Science and Technology*, *29*(6), 4929-4940.

Savary A., & Waszczuk J. (2020). Polish corpus of verbal multiword expressions. In *Proceedings of the Joint Workshop on Multiword Expressions and Electronic Lexicons*, (pp. 32–43). Barcelona, Spain: Association for Computational Linguistics.

Schindler, M., & Domahidi, E. (2021). The growing field of interdisciplinary research on user comments: A computational scoping review. *New Media & Society, 23*(8). doi: 10.1177/1461444821994491

Shahrour, I., & Xie X. (2021). Role of Internet of Things (IoT) and Crowdsourcing in Smart City Projects. *Smart Cities*, *4*(4), 1276-1292. doi: 10.3390/smartcities4040068

Stępniak, C., Jelonek, D., Wyrwicka, M., & Chomiak-Orsa, I. (2021). Integration of the Infrastructure of Systems Used in Smart Cities for the Planning of Transport and Communication Systems in Cities. *Energies*, *14*(11), 3069. doi: 10.3390/en14113069

Szumski, O. (2020). Technological trust from the perspective of digital payment. *Procedia Computer Science, 176*, 3545-3554. doi: 0.1016/j.procs.2020.09.032

Szymkowiak, M., Klimanek, T., & Józefowski, T. (2018). Applying Market Basket Analysis to Official Statistical Data. *Econometrics*, *22*(1), 39-57. doi: 10.1261/rna.7252205

Vermeeren, A., Law E., Roto, V., Obrist, M., Hoonhout, J., & Väänänen, K. (2010). User experience evaluation methods: current state and development needs. In *Proceedings of the 6th Nordic Conference on Human-Computer Interaction: Extending Boundaries (NordiCHI '10)* (pp. 521–530). New York, USA: Association for Computing Machinery.

Wieczorkowski, J. (2019). Open Data as a Source of Product and Organizational Innovations, In *Proceedings of the 14th European Conference on Innovation and Entrepreneurship ECIE 2019*, *2* (pp. 1118–1127).

Wieczorkowski, J., Chomiak-Orsa, I., & Pawełoszek, I. (2021). *Big data w zarządzaniu* [*Big Data in Management*], Warszawa, Poland: Polskie Wydawnictwo Ekonomiczne.

World Bank (2022). *Urban Development.* Retrieved from https://www.worldbank.org/en/topic/urbandevelopment/overview

Zavitsas, K., Kaparias, I., Bell, M., & Tomassini, M. (2010). Transport problems in cities. *ISIS*, *1*. Retrieved from https://trimis.ec.europa.eu/sites/default/files/project/documents/20120402_173932_45110_D%201.1%20-%20Transport%20problems%20in%20cities%20-%20v3.pdf

Zhuge, H., Zheng, L., Zhang, N., & Li, X. (2004). An automatic semantic relationships discovery approach. In *Proceedings of the 2004 International Conference on the World Wide Web* (pp. 278-279). doi: 10.1145/1013367.1013434