Association Rule System for Effective Risk Management of a Cinema Chain

Andrii Roskladka[1], Nataliia Roskladka[1], Volodymyr Kulazhenko[1], Alona Desiatko[1], and Alona Shtanova[1]

[1] *State University of Trade and Economics, 19 Kyoto str., Kyiv, 02156, Ukraine*

Abstract
The object of research is risk management in the activity of a cinema chain. The solution to the tasks set in the research requires a scientific systematic approach using modern data analysis technologies. It can ensure the formation of an effective strategy for the development of the cinema chain business, making optimal management decisions, and rapid profit growth. The main groups of services, that cinema guests used during their visit, were determined by expert methods. Based on the responses of about a thousand cinema-goers, patterns in their purchasing behavior were identified. To analyze the risk management process of a cinema chain, the authors propose using data association methods. The procedure for generating popular sets of services for cinema goers and forming association rules for choosing services have been carefully investigated. The qualitative characteristics of association rules were calculated and ranked according to various criteria. The R and Python software tools for forming association rules were analyzed. The interpretation of "strong" rules for a clear understanding of the organization of cinema activities under minimal risks and the construction of its communication policy was presented.

Keywords
Risk management, cinema chain, data association, association rules, apriori algorithm.

1. Introduction

The modern process of forming effective risk management is based on understanding the needs of customers and the ability to accurately predict their behavior. In the process of forming the risk management system of a cinema chain, researchers apply general scientific and analytical methods that help to increase its effectiveness. The most relevant source of information on which the risk management can be based is the existing consumer's observation because all aspects of the risk management are aimed at them. An in-depth research of consumer behavior will help to see in practice opportunities for more effective communication, product, and price components in the overall risk management strategy [1–3].

The sector of entertainment and recreation, which includes cinemas, has suffered perhaps the most in recent years. The difficult situation in the country causes significant risks in the activity of a cinema chain. All cinemas in Ukraine have been closed since the beginning of the COVID-19 pandemic since the end of March 2020. According to the information and analytical publication Media Business Reports [4], the negative impact of external conditions is reflected in the annual indicators of the box office in Ukraine. So, compared to 2018, the annual box office of 2019 increased by 20%, and in 2020 it decreased by 66%. This led to the forced downtime of cinemas and subsequent mass layoffs of their employees. During the next two years, during short-term periods of easing quarantine restrictions, the situation in the field of entertainment and

CPITS-2024: Cybersecurity Providing in Information and Telecommunication Systems, February 28, 2024, Kyiv, Ukraine
EMAIL a.roskladka@knute.edu.ua (A. Roskladka); n.roskladka@knute.edu.ua (N. Roskladka); v.kulazhenko@knute.edu.ua (V. Kulazhenko); desyatko@knute.edu.ua (A. Desiatko); a.shtanova@knute.edu.ua (A. Shtanova)
ORCID: 0000-0002-1297-377X (A. Roskladka); 0000-0001-7333-4050 (N. Roskladka); 0000-0002-3535-3442; (V. Kulazhenko); 0000-0003-2860-2188 (A. Desiatko); 0000-0003-1582-5219 (A. Shtanova)

recreation did not significantly improve. Those cinema chains that could not adapt to the new conditions were closed. Others were able to invent new ways of interacting with visitors and gradually restore and increase the pace of ticket sales.

In February 2022, with the beginning of the full-scale aggression of the Russian Federation in Ukraine, cinemas were closed again. The gradual opening of cinemas and the very slow recovery of business in this area began in the summer of 2022. However, it is still too early to talk about the improvement of the situation, because Ukraine is at war. A curfew is in effect throughout Ukraine, which forces cinemas to abandon most of the evening showings, which brought the greatest profit. The difficult situation forces us to reconsider the old and proven approaches to management in the field of providing film services. Therefore, it is necessary to apply new effective methods that are based on a systematic scientific study of user preferences and intelligent analysis of the results of these researches and can quickly and effectively affect the business performance of the cinema chains [5, 6].

2. Research Methodology

The theoretical and methodological justification of research is the fundamental principles of risk management and strategic analysis, a systematic approach, data analysis and synthesis, and a dialectical method in justifying the use of information technologies.

In particular, the following scientific methods are used in research:

- Risk analysis is used to identify unresolved problems in the management policy of cinemas and to develop methods for their elimination.
- Strategic analysis is used to identify unresolved problems in the strategic development of the cinema chain, and to justify the need to use data science methods for intelligent analysis of the strategy of the cinema chain’s activities.
- Association method is used to create cause-and-effect relationships between rating categories that were offered to cinema visitors.
- Apriori algorithm is used to form sets of popular sets of evaluation categories and to construct a system of association rules based on these sets.
- Graphic method is used to construct a scheme for forming subsets of different capacities from popular sets of evaluation categories.
- Method of quantitative analysis is used to calculate the characteristics of support and confidence of association rules, and their ranking and to identify a system of “strong” rules with the least risk in the activity of a cinema chain.

The informational basis for the research is data from a survey of respondents from different regions of Ukraine and different cinema chains.

3. Literature and Hypothesis Development

Risk management research in the field of picture motion industry is becoming the central topic of many studies because this type of business has its specifics, which is expressed in the need to sell its services both online and offline. Therefore, great attention should be paid both to Internet communications and to the creation of a special cinema atmosphere and service. Kim In Kyung [7], using cross-sectional data on Korean cinema chains, found evidence that productivity is higher in company-owned cinemas than in cinemas with a franchise. E. Salvador, J.-P. Simon, P.-J. Benghozi [8] explored the implementation of cinema value creation through innovation. B. Sheremeta, N. Chukhray, O. Karyy [9, 10]. studied the visitors to cinemas in Ukraine. According to their research, visitors believe that cinemas should develop such services as a cinema market, the establishment of an online cinema, the ability to leave a child with a nanny, special cinema uses, and a monthly subscription. Quite a lot of attention is paid to the impact of COVID-19 on the film distribution sphere [11].

Active scientific research in the field of association analysis began at the end of the twentieth century. Modern researchers often rely on the work “Fast Discovery of Association Rules” [12]. It is also necessary to note the works of R. Agrawal, R. Srikant [13], C. Borgelt [14, 15], and M. J. Zaki [16, 17].

The development of tools for finding association rules is also developing rapidly. It should be noted the scientific activity of M. Hashler, who covers the application of association rules in the R-package ecosystem [18–21].

The practical implementation of association rules is now actively used in many fields, such as economy [22, 23], tourism [24], medicine (especially in research related to COVID-19) [25, 26], education [27], etc. The development of the IT industry led to the application of association rules in text mining, in particular, in the analysis of messages in the social networks Facebook and Twitter [28, 29].

There are few examples of publications on association rules in the field of film industry [30, 31], but they relate to a rating of films and the construction of recommender systems for their selection.

The authors of the current research fully agree with O. Araz, T. Choi, D. Olson, and F. Salman [32] regarding the importance of the interaction between risk management and data science. Risk management and data science are large and independent sciences, but they cannot exist separately from each other. Risk management without data science is devoid of modern and effective data analysis technologies. Data science without risk management has no applied nature and loses its meaning without its practical application. Therefore, to obtain a synergistic effect from the use of data mining in risk management, an organic combination of advanced concepts of risk analysis and modern data mining technologies and tools is required.

4. Objective and Context of Research

The purpose of the scientific research is to replace inefficient situational management with a systematic approach in the risk management analysis of cinema chain activities, to ensure effective management decisions in the field of strategic development of the cinema chain based on marketing research and data mining obtained during the survey of cinema visitors.

Systematization of the complex structure of large volumes of data has led to the emergence of affinity analysis, which is one of the most common methods of data mining. Its purpose is to search for association rules to study the mutual connection between events that occur together.

For the first time, the problem of finding association rules was proposed for finding typical patterns of purchases made in supermarkets, so sometimes affinity analysis is called market basket analysis. Today, association rules are widely used not only in the retail sector but also in other areas, especially in the service sector.

The authors of the paper set a goal for themselves to apply the idea of affinity analysis and association rules for optimizing the management of cinema chain activities.

5. Creating a System of Association Rules

5.1. Structure of Initial Data for Forming Association Rules

The authors developed an online survey form for cinema chain visitors. It was decided to place an invitation to the survey directly in front of the entrance to the cinema hall, because viewers can buy tickets online and not use other cinema services. The invitation is a small postcard with brief information about the survey and a QR code that you can use to open the online form and complete the survey.

Viewers were asked the following questions and possible answers (the answer symbol is indicated in brackets):

1. How many tickets do you usually buy in one cinema visit?
   - one ticket *(A)*
   - two or more tickets *(B)*.
2. How do you most often buy cinema tickets:
   - online via the website or mobile app *(C)*.
   - at the cinema box office *(D)*.
3. What type of seat do you most often choose when buying tickets?
   - regular seats *(E)*
   - VIP seats *(F)*.
4. What days of the week do you most often visit the cinema?
   - on weekends *(G)*
   - on weekdays *(H)*.

5. At what time do you most often visit the cinema?
   - in the morning *(I)*
   - in the afternoon *(J)*
   - in the evening *(K)*
   - late in the evening *(L)*.
6. Do you prefer films in IMAX format? *(M)*
7. Do you prefer films in IMAX format? *(M)*
8. Do you usually buy products at the cinema bar before the film starts? *(O)*
9. Do you use your cinema's loyalty program? *(P)*
10. Are you a subscriber and an active user of your cinema's social network? *(Q)*

The survey was designed in such a way that viewers had to indicate only 1 answer to questions 1–5. Questions 6–10 had to be answered "yes" or "no".

947 respondents, who represented 8 regions of Ukraine, were visitors of 6 cinema chains and were aged from 18 to 60, were interviewed during 4 months (from October 2022 to January 2023).

For further processing, the respondents' answers were presented in the binary system, where "0" means that the respondent did not choose this answer to questions 1–5 or gave a negative answer to questions 6–10. Accordingly, "1" means that the respondent chose this answer to questions 1–5 or a positive answer to questions 6–10.

Table 1 presents a fragment of the initial data table, the total number and percentage of responses provided.

Table 1
Fragment of the Table of Initial Data, Presented in Binary Format

ID	*A*	*B*	*C*	*D*	*E*	*F*	*G*	*H*	*I*	*J*	*K*	*L*	*M*	*N*	*O*	*P*	*Q*
1	1	0	1	0	1	0	0	1	0	0	1	0	0	0	0	0	0
2	1	0	1	0	1	0	1	0	0	0	1	0	0	0	1	0	0
3	1	0	1	0	1	0	0	1	0	0	1	0	0	0	0	0	0
4	1	0	1	0	1	0	0	1	0	0	1	0	0	0	0	0	0
5	1	0	1	0	1	0	0	1	0	0	1	0	0	0	0	0	0
6	1	0	1	0	1	0	1	0	0	0	1	0	0	0	0	1	0
7	1	0	1	0	1	0	1	0	0	0	1	0	0	0	0	0	0
8	1	0	1	0	1	0	0	1	0	0	1	0	0	0	1	1	0
9	1	0	1	0	1	0	0	1	0	0	1	0	0	0	1	0	1
10	1	0	1	0	1	0	0	1	0	0	1	0	0	0	0	0	0
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
945	0	1	0	1	1	0	0	1	0	1	0	0	0	0	0	1	1
946	0	1	0	1	0	1	0	1	0	0	1	0	1	0	1	0	0
947	0	1	1	0	1	0	1	0	0	0	1	0	0	0	1	1	0
	352	595	691	256	614	333	271	676	31	107	660	149	118	11	720	419	128
	37%	63%	73%	27%	65%	35%	29%	71%	3%	11%	70%	16%	12%	1%	76%	44%	14%

5.2. Formation of a Set of Popular Itemsets

Different methods and algorithms of association analysis are used to form association rules. One of the most widespread algorithms that allows reducing the search space to sizes that provide acceptable computational and time costs is the Apriori algorithm [33]. This algorithm is based on the concept of a transaction as some set of compatible events and the concept of popular itemsets that are often found in different transactions.

To form sets of popular itemsets, it is necessary to set a minimum support threshold for the popularity of the itemset. The choice of such a threshold value significantly affects the number of rules. Too large a value of the minimum support threshold will dramatically reduce the number or even make it impossible to form association rules. On the contrary, too small a value will lead to the formation of a large number of rules with low reliability, and this will complicate their interpretation and practical implementation. Thus, the value of the minimum support threshold is set exclusively experimentally. The authors of this study set the value of the minimum support threshold of association rules at the level of 30%.

Therefore, the set of popular 1-itemsets will consist of criteria that have support no less than 30%:

$$F_1 = \{A, B, C, E, F, H, K, O, P\} \quad (1)$$

Next, we find popular 2-itemsets, forming all 36 possible combinations of two elements from *F*$_1$: *AB, AC, AE, AF, AH, AK, AO, AP, BC, BE. BF, BH, BK, BO, BP, CE, CF, CH, CK, CO, CP, EF, EH, EK, EO, EP, FH, FK, FO, FP, HK, HO, HP, KO, KP, OP.*

To estimate the level of support of a 2-itemset, we need to multiply the corresponding values from Table 1. For example, for respondent ID = 1, the value of the itemset AB = $1 \cdot 0 = 0$, and the value of the itemset AC = $1 \cdot 1 = 1$. A fragment of the table with support level calculations for each 2-itemset is presented in Table 2.

Table 2

A Fragment of the Table for Calculating the Level of Support of 2-itemsets

ID	*AB*	*AC*	*AE*	*AF*	*AH*	*AK*	*AO*	*AP*	*BC*	*BE*	*BF*	*BH*	*BK*	*BO*	*...*	*KP*	*OP*
1	0	1	1	0	1	1	0	0	0	0	0	0	0	0	...	0	0
2	0	1	1	0	0	1	1	0	0	0	0	0	0	0	...	0	0
3	0	1	1	0	1	1	0	0	0	0	0	0	0	0	...	0	0
4	0	1	1	0	1	1	0	0	0	0	0	0	0	0	...	0	0
5	0	1	1	0	1	1	0	0	0	0	0	0	0	0	...	0	0
6	0	1	1	0	0	1	0	1	0	0	0	0	0	0	...	1	0
7	0	1	1	0	0	1	0	0	0	0	0	0	0	0	...	0	0
8	0	1	1	0	1	1	1	1	0	0	0	0	0	0	...	1	1
9	0	1	1	0	1	1	1	0	0	0	0	0	0	0	...	0	0
10	0	1	1	0	1	1	0	0	0	0	0	0	0	0	...	0	0
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
945	0	0	0	0	0	0	0	0	0	1	0	0	0	0	...	0	0
946	0	0	0	0	0	0	0	0	0	0	0	0	0	0	...	0	0
947	0	0	0	0	0	0	0	0	1	1	0	0	0	0	...	1	1
	0	278	294	58	242	263	249	164	413	320	275	434	397	471	...	274	324
	0%	29%	31%	6%	26%	28%	26%	17%	44%	34%	29%	46%	42%	50%	...	29%	34%

Naturally, 2-itemsets *AB* and *EF* have a support level of 0%, because the answers mutually exclude each other. Out of 36 pairs of values, 19 itemsets have support with a minimum level of 30%, which forms a set of popular itemsets.

$$F_2=\{AE, BC, BE, BH, BK, BO, CE, CH, CK, CO, CP, EH, EK, EO, FO, HK, HO, KO, OP\} \quad (2)$$

Using popular itemsets F2, we generate sets of 3-itemsets F3. To do this, we need to connect the set F2 to itself by choosing connecting itemsets. *k*-itemsets are connected if they have *k*-1 common items. For example:

$$\{AE\}+\{BE\}=\{ABE\} \quad (3)$$

Generating itemsets based on simple combinations of elements of the previous level quickly leads to an uncontrolled increase in unpromising combinations. Various analytical [34] and algorithmic [35] methods of combinatorial optimization can be used to prevent such a negative situation. The practical implementation of such approaches to finding associative rules in the Apriori algorithm is the property of antimonotonicity: if *N* is not a popular itemset, then adding some new item *M* to the itemset *N* does not make it more popular, i. e. the itemset $M \cup N$ will not be popular. For example, in our research, the combination of the popular *AE* and *BE* itemsets results in the *ABE* 3-item set. But it won't be popular because it includes *AB* 2-itemset that has zero support. Adding element *E* to its composition will not increase the support of this itemset and it will remain at the level of 0%.

The set of all possible 40 connected itemsets will be *ABE, ACE, AEH, AEK, AEO, BCE, BCH, BCK, BCO, BCP, BEH, BEK, BEO, BHK, BHO, BFO, BKO, BOP, CEH, CEK, CEO, CEP, CHK, CHO, CHP, CKO, CKP, COP, CFO, EHK, EHO, EKO, EFO, EOP, FHO, FKO, FOP, HKO, HOP, KOP*.

The property of antimonotonicity made it possible to exclude 19 unpromising itemsets from this list and to leave 21 itemsets for further verification: *BCE, BCH, BCK, BCO, BEH,*

BEK, BEO, BHK, BHO, BKO, CEH, CEK, CEO, CHK, CHO, CKO, COP, EHK, EHO, EKO, HKO.

Table 3 presents calculations of the level of support for 3-itemsets.

16 itemsets have support with a minimum level of 30%, of the 21 3-itemsets of values pairs, which form a set of popular itemsets:

Table 3

A Fragment of the Table for Calculating the Support Level of 3-itemsets

ID	*BCE*	*BCH*	*BCK*	*BCO*	*BEH*	*BEK*	*BEO*	*BHK*	*BHO*	*BKO*	*CEH*	*CEK*	*...*	*EKO*	*HKO*
1	0	0	0	0	0	0	0	0	0	0	1	1	...	0	0
2	0	0	0	0	0	0	0	0	0	0	0	1	...	1	0
3	0	0	0	0	0	0	0	0	0	0	1	1	...	0	0
4	0	0	0	0	0	0	0	0	0	0	1	1	...	0	0
5	0	0	0	0	0	0	0	0	0	0	1	1	...	0	0
6	0	0	0	0	0	0	0	0	0	0	0	1	...	0	0
7	0	0	0	0	0	0	0	0	0	0	0	1	...	0	0
8	0	0	0	0	0	0	0	0	0	0	1	1	...	1	1
9	0	0	0	0	0	0	0	0	0	0	1	1	...	1	1
10	0	0	0	0	0	0	0	0	0	0	1	1	...	0	0
...	...	...	...	...	...	...	...	...	...	...	...		...	...	...
945	0	0	0	0	1	0	0	0	0	0	0	0	...	0	0
946	0	0	0	0	0	0	0	1	1	1	0	0	...	0	1
947	1	0	1	1	0	1	1	0	0	1	0	1	...	1	0
	196	297	288	331	232	197	235	294	344	339	295	299	...	315	388
	21%	31%	30%	35%	24%	21%	25%	31%	36%	36%	31%	32%	...	33%	41%

The process of forming sets of popular itemsets F_k continues as long as the elements of such set contain connecting itemsets with *k-1* common elements. Let's try to form a set of 4-itemsets. We have the following result of forming from 14 elements: *BCHK, BCHO, BCKO, BCEK, BCEO, BHKO, BEHK, BEHO, BEKO, CEHK, CEHO, CEKO, CHKO, EHKO.*

Applying the property of antimonotonicity, we determine unpromising elements: *BCEK, BCEO* (contain nonpopular subset *BCE), BEHK, BEHO* (contain nonpopular subset *BEH), BEKO* (contain nonpopular subset *BEK*). We will check the remaining 9 4-itemsets for popularity by calculating the corresponding supports (Table 4).

Table 4

A Fragment of the Table for Calculating the Support Level of 4-itemsets

ID	*BCHK*	*BCHO*	*BCKO*	*BHKO*	*CEHK*	*CEHO*	*CEKO*	*CHKO*	*EHKO*
1	0	0	0	0	1	0	0	0	0
2	0	0	0	0	0	0	1	0	0
3	0	0	0	0	1	0	0	0	0
4	0	0	0	0	1	0	0	0	0
5	0	0	0	0	1	0	0	0	0
6	0	0	0	0	0	0	0	0	0
7	0	0	0	0	0	0	0	0	0
8	0	0	0	0	1	1	1	1	1
9	0	0	0	0	1	1	1	1	1
10	0	0	0	0	1	0	0	0	0
...	...	...	...	...	...	...	...	...	...
945	0	0	0	0	0	0	0	0	0
946	0	0	0	1	0	0	0	0	0
947	0	0	1	0	0	0	1	0	0
	206	237	246	251	206	217	226	283	231
	22%	25%	26%	27%	22%	23%	24%	30%	24%

Only one element satisfies the minimum support condition. Thus, the set

$$F_4 = \{CHKO\} \quad (4)$$

It is impossible to form connecting itemsets from one element, so the first stage of the Apriori algorithm (the process of forming popular itemsets) is completed. It has the following results in our research:

$$F_1 = \{A, B, C, E, F, H, K, O, P\} \quad (5)$$

$$F_2 = \{AE, BC, BE, BH, BK, BO, CE, CH, CK, CO, CP, EK, EO, FO, HK, HO, KO, OP\} \quad (6)$$

$$F_3 = \{BCH, BCK, BCO, BHK, BHO, BKO, CEH, CEK, CEO, CHK, CHO, CKO, EHK, EHO, EKO, HKO\} \quad (7)$$

$$F_4 = \{CHKO\} \quad (8)$$

The graphic model of the formation of sets F_1, F_2, F_3, F_4 is presented in Fig. 1.

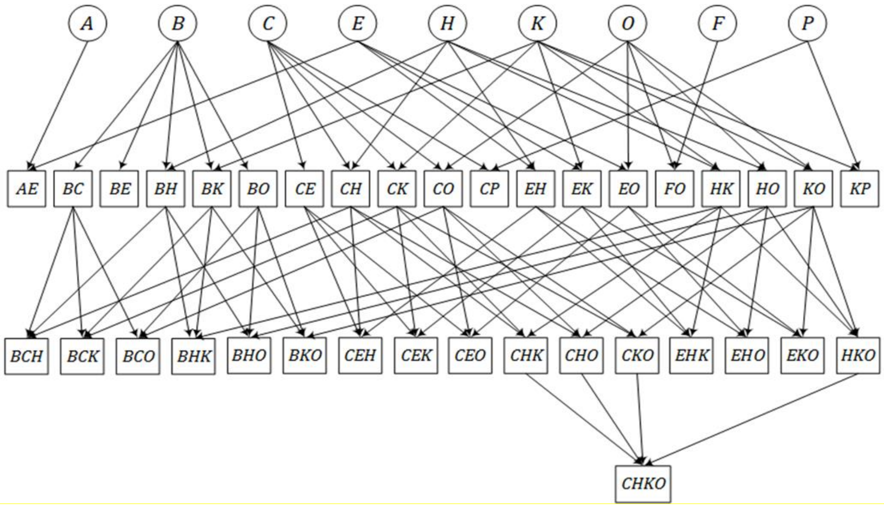


Figure 1: A graphic model of the formation of sets of popular itemsets of different capacities

5.3. Formation of Association Rules

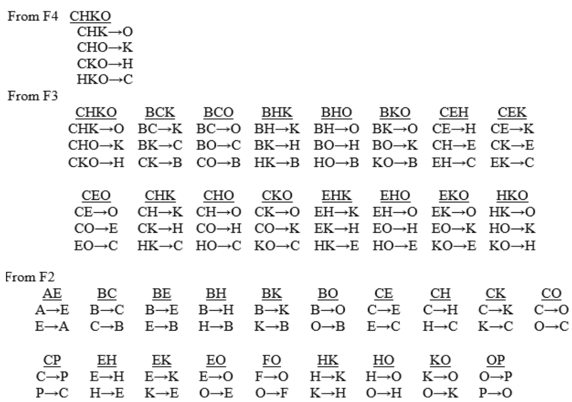
The second step of the Apriori algorithm is the direct formation of association rules. In general, the Apriori algorithm does not impose restrictions on the number of elements in both parts of the association rule. However, in practice, only rules that contain one conclusion are often considered. At the same time, the number of components of the condition is not limited. This approach pursues two goals:

1. The number of rules is reduced, which makes the calculation easier;
2. The presence of one item in the conclusion makes the rule unambiguous for interpretation, and ready for practical implementation.

According to this approach, in each set of popular itemsets $F_i, i \geq 2$, it is necessary to select such itemsets that contain i items. Then you should also select *i* one-element subsets from each such itemset. The obtained subsets will form the conclusions of the association rule, and the rest *i-1* elements will be combined into appropriate conditions by conjunction.

For example, two rules (*A→E, E→A*) can be formed from the popular itemset *AE* belonging to the set F_2; three rules (*BH→K, BK→H, HK→B)* can be formed from the itemset *BHK* belonging to the set F_3; four rules *(CHK→O, CHO→K, CKO→H, HKO→C)* can be formed from a itemset of CHKO belonging to the set F_4. The procedure is repeated until all popular itemsets are exhausted.

In the task of forming the risk management strategy for a cinema chain, we will have the following set, which consists of 90 association rules (Fig. 2):


From F4

CHKO
CHK→O
CHO→K
CKO→H
HKO→C

From F3

CHKO	BCK	BCO	BHK	BHO	BKO	CEH	CEK
CHK→O	BC→K	BC→O	BH→K	BH→O	BK→O	CE→H	CE→K
CHO→K	BK→C	BO→C	BK→H	BO→H	BO→K	CH→E	CK→E
CKO→H	CK→B	CO→B	HK→B	HO→B	KO→B	EH→C	EK→C

CEO	CHK	CHO	CKO	EHK	EHO	EKO	HKO
CE→O	CH→K	CH→O	CK→O	EH→K	EH→O	EK→O	HK→O
CO→E	CK→H	CO→H	CO→K	EK→H	EO→H	EO→K	HO→K
EO→C	HK→C	HO→C	KO→C	HK→E	HO→E	KO→E	KO→H

From F2

AE	BC	BE	BH	BK	BO	CE	CH	CK	CO
A→E	B→C	B→E	B→H	B→K	B→O	C→E	C→H	C→K	C→O
E→A	C→B	E→B	H→B	K→B	O→B	E→C	H→C	K→C	O→C

CP	EH	EK	EO	FO	HK	HO	KO	OP
C→P	E→H	E→K	E→O	F→O	H→K	H→O	K→O	O→P
P→C	H→E	K→E	O→E	O→F	K→H	O→H	O→K	P→O

Figure 2: An association rules of cinema chain's set

5.4. Quantitative Characteristics of Association Rules

Association rules describe the relationship between itemsets, which are characterized by two main indicators, they are *support S* and *confidence C*.

Support of association rule is the element of transactions that contain as a condition, as a consequence. For example, for a rule *A→B* support *S(A→B)* means the ratio of the number of transactions *AB*, that simultaneously contain condition *A* and consequence *B*, to the total number of transactions.

The confidence *C(A→B)* of an association rule *A→B* is defined as the ratio of the number of transactions, that simultaneously contain condition *A* and consequence *B*, to the number of transactions that contain only condition *A*.

The support and confidence of association rules (or their product) are most often used to rank the obtained rules in descending order of importance and highlight a subset of so-called "strong" rules, that have a minimal level of risk. Their interpretation and consideration will bring the most tangible effect to the business. In addition to these characteristics, *L (lift), T (leverage),* and *I (improvement)* are also used for qualitative assessment of association rules [33].

Table 5 presents a set of obtained association rules with calculated quantitative characteristics.

Table 5
A Set of Association Rules with Support and Confidence Characteristics

ID	Rule	S	C	ID	Rule	S	C	ID	Rule	S	C
1	CHK→O	0,300	0,82	31	EO→C	0,315	0,683	61	B→K	0,419	0,667
2	CHO→K	0,300	0,757	32	CH→K	0,364	0,713	62	K→B	0,419	0,602
3	CKO→H	0,300	0,702	33	CK→H	0,364	0,686	63	B→O	0,497	0,792
4	HKO→C	0,300	0,729	34	HK→C	0,364	0,732	64	O→B	0,497	0,654
5	BC→H	0,314	0,719	35	CH→O	0,395	0,773	65	C→E	0,442	0,606
6	BH→C	0,314	0,684	36	CO→H	0,395	0,71	66	E→C	0,442	0,682
7	CH→B	0,314	0,614	37	HO→C	0,395	0,723	67	C→H	0,511	0,7
8	BC→K	0,304	0,697	38	CK→O	0,426	0,801	68	H→C	0,511	0,716
9	BK→C	0,304	0,725	39	CO→K	0,426	0,765	69	C→K	0,531	0,728
10	CK→B	0,304	0,573	40	KO→C	0,426	0,756	70	K→C	0,531	0,762
11	BC→O	0,35	0,801	41	EH→K	0,309	0,674	71	C→O	0,556	0,763
12	BO→C	0,35	0,703	42	EK→H	0,309	0,711	72	O→C	0,556	0,732
13	CO→B	0,35	0,628	43	HK→E	0,309	0,622	73	C→P	0,364	0,499
14	BH→K	0,31	0,677	44	EH→O	0,332	0,722	74	P→C	0,364	0,823
15	BK→H	0,31	0,741	45	EO→H	0,332	0,72	75	E→H	0,459	0,708
16	HK→B	0,31	0,624	46	HO→E	0,332	0,607	76	H→E	0,459	0,643
17	BH→O	0,363	0,793	47	EK→O	0,333	0,765	77	E→K	0,435	0,671
18	BO→H	0,363	0,73	48	EO→K	0,333	0,722	78	K→E	0,435	0,624
19	HO→B	0,363	0,665	49	KO→E	0,333	0,591	79	E→O	0,46	0,71
20	BK→O	0,358	0,854	50	HK→O	0,41	0,824	80	O→E	0,46	0,606
21	BO→K	0,358	0,72	51	HO→K	0,41	0,75	81	F→O	0,3	0,853
22	KO→B	0,358	0,636	52	KO→H	0,41	0,728	82	O→F	0,3	0,394
23	CE→H	0,312	0,704	53	A→E	0,31	0,835	83	H→K	0,497	0,697
24	CH→E	0,312	0,61	54	E→A	0,31	0,479	84	K→H	0,497	0,714
25	EH→C	0,312	0,678	55	B→C	0,436	0,694	85	H→O	0,546	0,765
26	CE→K	0,316	0,714	56	C→B	0,436	0,598	86	O→H	0,546	0,718
27	CK→E	0,316	0,594	57	B→E	0,338	0,538	87	K→O	0,563	0,808
28	EK→C	0,316	0,726	58	E→B	0,338	0,521	88	O→K	0,563	0,74
29	CE→O	0,315	0,711	59	B→H	0,458	0,729	89	O→P	0,342	0,45
30	CO→E	0,315	0,565	60	H→B	0,458	0,642	90	P→O	0,342	0,773

Thus, as a result of the application of the Apriori algorithm, it was possible to identify 90 associative rules that show which services from the initial set of transactions will most often be chosen by cinemagoers together. Such indicators will make it possible to build logical connections between various services and to develop a perfect risk management strategy in the activity of a cinema chain.

5.5. Interpretation of Association Rules

From a practical point of view, the interpretation of the obtained association rules is important. Let's consider examples of the interpretation of several "strong" rules, ordered in Table 6 by the level of confidence.

ID 1: CHK→O. A visitor who buys a ticket online for a weekday evening show will buy the products at the bar. The rule has a minimum allowable support of 30%. This means that 30% of the surveyed cinema-goers indicated that they use all these four services. However, this rule has one of the highest confidence values—82%, which confirms its high significance. From a strategic point of view, this rule distinguishes with high reliability the customers who will be served at the bar. Therefore, the bar must be opened on weekdays in the evening. It is possible to limit its functioning to other days and hours if the number of bartenders is insufficient. It is also useful to allow ordering online bar products for evening showings.

Table 6
A Set of "Strong" Rules by Level of Confidence

ID	Rule	C
20	BK→O	0,854
81	F→O	0,853
53	A→E	0,835
50	HK→O	0,824
74	P→C	0,823
1	CHK→O	0,820
87	K→O	0,808
11	BC→O	0,801
38	CK→O	0,801
17	BH→O	0,793

ID 53: A→E. A visitor who buys one ticket chooses a regular seat. Rule support is 31%, and confidence is 83.5%. This rule demonstrates the fact that it is important not only to use the rule directly but also to try to change it. The cinema is interested in buying VIP seats, which provide more profit. Therefore, management should take measures to encourage visitors to choose VIP seats. This can be achieved by introducing a broader loyalty program for visitors with VIP seats, for example, discounts on certain products in the cinema bar.

ID 74: P→C. Most visitors, who use the loyalty program, buy tickets online. 36.4% of surveyed visitors are members of the loyalty program, and 82.3% of such participants buy tickets directly through the cinema chain's website or a mobile application. It is necessary to expand the loyalty program in the development of the risk management strategy and encourage participation in it in every possible way to get an increase in online ticket purchases. It is necessary to increase its accessibility and clarity for people of different ages and to conduct the communication policy of the cinema through it. That is, members of the loyalty program should receive e-mails, push notifications from the mobile apps, newsletters on messengers for birthdays, on the occasion of the release of new films, as well as receive special offers, they are promotional codes with discounts on tickets and bar products, invitations to special parties and meetings with actors, etc.

5.6. Software Tools for Forming Association Rules

In the case of a large number of transactions and criteria in each of them, it is appropriate to use software tools for generating and processing association rules.

For the practical application of association rules, you need to use software environments, in which data mining algorithms are implemented. Let's briefly describe the process of processing association rules in the most advanced data analytics environments, such as *R* and *Python*.

To implement the *Apriori* algorithm in *R* [19], you need to install the *arules* library and to visualize association rules, you need to install the *arulesViz* library:

```
# Installing Packages
install.packages("arules")
install.packages("arulesViz")
 # Loading package
library(arules)
library(arulesViz)
```

Formation of the system of association rules is performed using the *apriori* function:

```
apriori (data = dataset, parameters)
```

The parameters of association rules are most often set as a list with the minimum values of the numerical characteristics of the rules. For example,

rules = apriori (data = dataset, parameter = list(support = 0.004, confidence = 0.2))

A bar chart of the relative frequency of transaction elements is often used as a visualization:

 # Plot
itemFrequencyPlot(dataset, topN = 10)

To visualize the resulting system of rules, you should sort the rules by one of the quantitative characteristics, and then build a model for a certain number of "strong" rules. In the example below, the system of 10 "strong" rules is sorted by the lift characteristic:

 # Visualizing the results
inspect(sort(rules, by = 'lift')[1:10])
plot(rules, method = "graph", measure = "confidence", shading = "lift")

The *Python* ecosystem also provides several tools for finding association rules in datasets. *Python* modules for data analysis, such as *pandas* and *numpy*, which are familiar for data analysis, are used to collect, store, manipulate, and prepare datasets. The implementation of the backend part of these modules is developed in the *C++* programming language, which makes it possible to combine the flexibility, inherent in *Python*, and the speed of performing large volumes of operations.

The most popular data analysis modules, such as *tensorflow* and *sklearn*, do not include algorithms for searching for association rules at the time of publication. So, you can find many original implementations of this tool on the web. The *mlxtend* module is most often used, which allows you to perform data pre-processing, generate sets of rules, and search for relevant dependencies [36, 37]. *Matplotlib* and *plotly* modules are used to visualize the obtained results. They make it possible to display any graphic information, as well as to build charts of any complexity. Both packages include the ability to build interactive visualizations. In addition, the *plotly* module can be integrated into Flask web applications as a full-fledged dashboard. The *NetworkX* module is used to display graphs and structured information, which simplifies their visualization.

6. Conclusions

The entertainment sector and, in particular, the sphere of cinema activity experienced significant shocks during the COVID-19 pandemic. In Ukraine, the difficult situation worsened significantly with the beginning of Russia's military aggression. This demonstrated the significant dependence of the entertainment sector on the political and social situation of the country, and the loss of stability of the market of film services.

The management of the cinema chain is forced to look for new approaches to business management to minimize risks in their activities. Improvements in data collection methods in combination with new intelligent processing technologies are capable of creating real progress in the strategic and risk management analysis of the activities of cinema chains.

In their research, the authors propose to use association rules as one of the advanced data mining technologies to improve the risk management strategy of the cinema chain. Association rules offer two undeniable advantages. The first is that association analysis allows you to discover hidden logical connections and patterns that are unrealistic to track for a large amount of data. The second is that the results of the analysis in the form of association rules are ready for use. They indicate the directions for improvement of strategic planning and risk management policy, based on the synergy of subjective and objective factors, and also identify weaknesses that the cinema administration should pay attention to.

Thus, the application of a system of association rules for the formation of the risk management strategy has praxeological value for implementation in the daily work of a cinema chain.

References

[1] H. Shevchenko, et al., Information Security Risk Analysis SWOT, in: Workshop on Cybersecurity Providing in

Information and Telecommunication Systems, CPITS, vol. 2923 (2021) 309–317.
[2] D. Berestov, et al., Analysis of Features and prospects of Application of Dynamic Iterative Assessment of Information Security Risks, in: Workshop on Cybersecurity Providing in Information and Telecommunication Systems, CPITS, vol. 2923 (2021) 329–335.
[3] S. Shevchenko, et al., Information Security Risk Management using Cognitive Modeling, in: Workshop on Cybersecurity Providing in Information and Telecommunication Systems II, CPITS-II, vol. 3550 (2023) 297–305.
[4] Media Business Reports (2020). URL: https://mbr.com.ua.
[5] F. Kipchuk, et al., Assessing Approaches of IT Infrastructure Audit, in: IEEE 8th International Conference on Problems of Infocommunications, Science and Technology (2021). doi: 10.1109/picst54195.2021.9772181
[6] V. Buriachok, V. Sokolov, P. Skladannyi, Security Rating Metrics for Distributed Wireless Systems, in: Workshop of the 8th International Conference on "Mathematics. Information Technologies. Education": Modern Machine Learning Technologies and Data Science, vol. 2386 (2019) 222–23
[7] I. Kim, Organizational Form and Performance: The Cinema Chain Case, Appl. Econom. 53(56) (2021) 6472–6487. doi: 10.1080/00036846.2021.1946003.
[8] E. Salvador, J.-P. Simon, P.-J. Benghozi, Facing Disruption: The Cinema Value Chain in the Digital Age, Int. J. Arts Manag. 22(1) (2019) 25–40.
[9] B. Sheremeta, N. Chukhrai, Using the Blue Ocean Strategy by Ukrainian Cinema Networks in Uncertain Environment, Innov. Technol. Sci. Solut. Ind. 4(14) (2020) 137–146. doi: 10.30837/itssi.2020.14.137.
[10] B. Sheremeta, N. Chukhray, O. Karyy, Marketing Tools as the Competitiveness Enhancer of the Ukrainian Film Distribution Market Entities, Innov. Mark. 15(4) (2019) 88–101. doi: 10.21511/im.15(4).2019.08.
[11] S. Melnychenko, et al., Communication Policy of Cinema Industry Enterprises in the Context of COVID-19 (on the Example of Cinema Chains), Innov. Mark. 17(2) (2021) 112–124. doi: 10.21511/im.17(2).2021.11.
[12] R. Agrawal, et al., Fast Discovery of Association Rules, Advances in Knowledge Discovery and Data Mining 12(1) (1996) 307–328.
[13] R. Agrawal, R. Srikant, Fast Algorithms for Mining Association Rules, 20th Int. Conf. Very Large Data Bases (VLDB) 1215 (1994) 487–499.
[14] C. Borgelt, Efficient Implementations of Apriori and Eclat, FIMI'03: Proceedings of the IEEE ICDM Workshop on Frequent Itemset Mining Implementations (2003).
[15] C. Borgelt, Frequent Item Set Mining, WIREs: Data Mining and Knowledge Discovery 2(6) (2012) 437–456. doi: 10.1002/widm.1074.
[16] M. Zaki, Generating Non-Redundant Association Rules, 6th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (2000) 34–43. doi: 10.1145/347090.347101.
[17] M. Zaki, Scalable Algorithms for Association Mining, IEEE Transactions on Knowledge and Data Engineering 12(3) (2000) 372–390. doi: 10.1109/69.846291.
[18] M. Hahsler, et al., Associative Classification in R: arc, arulesCBA, and rCBA, The R J. 11(2) (2019) 254–267. doi: 10.32614/rj-2019-048.
[19] M. Hahsler, ArulesViz: Interactive Visualization of Association Rules with R, The R J. 9(2) (2017) 163–175. doi: 10.32614/RJ-2017-047.
[20] M. Hahsler, et al., The Arules R-package Ecosystem: Analyzing Interesting Patterns from large Transaction Data Sets, The J. Mach. Learn. Res. 12 (2011) 2021–2025. doi: 10.5555/1953048.2021064.
[21] M. Hahsler, B. Grün, K. Hornik, Arules-A Computational Environment for Mining Association Rules and Frequent Item Sets, J. Stat. Softw. 14 (2005) 1–25.
[22] Z. Huang, X. Lu, H. Duan, Mining Association Rules to Support Resource Allocation in Business Process Management, Expert Syst. Appl. 38(8)

(2011) 9483–9490. doi: 10.1016/j.eswa.2011.01.146.
[23] B. Öztayşi, P. Yurdadön, S. Onar, Fuzzy Clustering Based Association Rule Mining: A Case Study on Ecommerce, International Conference on Intelligent and Fuzzy Systems (2022) 112–118.
[24] D. Atsa'am, R. Wario, Association Rules on the COVID-19 Variants of Concern to Guide Choices of Tourism Destinations, Curr. Issues Tour. 25(10) (2022) 1536–1540. doi: 10.1080/13683500.2021.1951182.
[25] M. Tandan, et al., Discovering Symptom Patterns of COVID-19 Patients Using Association Rule Mining, Comput. Biol. Med. 131 (2021). doi: 10.1016/j.compbiomed.2021.104249.
[26] M. Wang, et al., The Application Characteristics of Traditional Chinese Medical Science Treatment on Vertigo Based on Data Mining Apriori Algorithm, Int. J. Wirel. Mob. Comput. 4 (2015) 349–354.
[27] A. Roskladka, et al., The Data Association Algorithm for the Formation of Optional IT-courses List System, in: ICT in Education, Research and Industrial Applications. Integration, Harmonization and Knowledge Transfer Vol. 2732 (2020) 515–531.
[28] M. Alanezi, Community Detection in Facebook Using Visual Approach and Clustering, J. Phys. Conf. Series 1804(1) (2021). doi: 10.1088/1742-6596/1804/1/012047.
[29] J. Diaz-Garcia, M. Ruiz, M. Martin-Bautista, Generalized Association Rules for Sentiment Analysis in Twitter, International Conference on Flexible Query Answering Systems, LNAI 11529 (2019) 166–175. doi: 10.1007/978-3-030-27629-4_17.
[30] Z. Zubi, A. Elrowayati, I. Abu Fanas, A Movie Recommendation System Design Using Association Rules Mining and Classification Techniques, WSEAS Transactions on Computers 21 (2022) 189–199. doi: 10.37394/23205.2022.21.24.
[31] K. Bury, Association Rule for Movie Ratings (2021). URL: https://rpubs.com/Bury/AssociationRuleForMovieRatings
[32] O. Araz, et al., Data Analytics for Operational Risk Management, Decision Sci. 51(6) (2020) 1316–1319. doi: 10.1111/deci.12443.
[33] D. Dietrich, B. Heller, B. Yang, Data Science & Big Data Analytics: Discovering, Analyzing, Visualizing and Presenting Data, John Wiley & Sons (2015).
[34] O. Emets, A. Roskladka, On Estimates of Minima of Criterion Functions in Optimization on Combinations, Ukrainian Math. J. 51(8) (1999) 1262–1265. doi: 10.1007/BF02592514.
[35] O. Yemets, A. Roskladka, Algorithmic Solution of Two Parametric Optimization Problems on a Set of Complete Combinations, Cybern. Syst. Anal. 35(6) (1999) 981–986. doi: 10.1007/BF02742292.
[36] Python Implementation of Apriori Algorithm (2023). URL: https://github.com/asaini/Apriori
[37] Advanced Python Visualization Library for Association Rules (2023). URL: https://github.com/Mazeofthemind/PyARMViz