UCBVis: Understanding Customer Behavior Sequences with Visual Interactive System

Md Rafiqul Islam*, Imran Razzak†, Xianzhi Wang‡, Peter Tilocca§, Guandong Xu*
*Advanced Analytics Institute (AAi), University of Technology Sydney, Australia
†School of Information Technology, Deakin University, Geelong, Australia
‡School of Computer Science, University of Technology Sydney, Australia
§OnePath Life Insurance, Australia
{mdrafiqul.islam-1@student, xianzhi.wang, guandong.xu}@uts.edu.au,
imran.razzak@deakin.edu.au, peter.tilocca@onepath.com.au

***Abstract*—Understanding customer behaviour (UCB) sequences with the multi-dimensional and temporal of the data is necessary for any competitive and global business aiming to provide interesting insights and to improve business strategies. While existing researchers have applied various data analytics approaches to understand and analyze behaviors of customer, they often failed to allow the analysts including business management, product marketing and development, and decision making, etc. Achieving these goals in collaboration with domain experts, we conducted a design study contributes to address a known problem with a novel solution and to provide data-driven visual decision support in collective policy data. We determine core study demands and then use a Visual Interactive System for Understanding Customer Behavior, named *UCBVis* that enables decision makers to gain detail insights into customer activities. In this study, we present customer behaviour pattern of multi-dimensional relationship through the visualisation system based on interweaving the pattern mining and querying with a designed encoding scheme. We use a large number of customer claim records and present visual outcomes to facilitate the exploration of customer behavior. Furthermore, we provide a concise set of insights and challenges associated with the use of *UCBVis* in the life insurance industry. We show the robustness of *UCBVis* through a user study with five participants shows that *UCBVis* is perceived to be more useful and provides actionable insights.**

***Index Terms*—Customer Behavior, Data analytics, Visualization, Behavioral pattern, Decision making.**

I. Introduction

The understanding of customer behavioral sequences is important in the life insurance industry in Australia aiming to provide a more insightful information's improving business operations and decision makings [1], [2]. This means insurance authorities such as insurance managers (IMs) can gain insights into customer such as how customer claimed at original state or suburb in Australia, and how frequently they claimed. However, because of the lack of proper understanding and computational expertise, IMs often cannot collect exact information of customer. Thus, it is becoming a severe economic problem [3]. For example, away from less interest for authority budgets, claim-cause evasion has contributed to a bias business race environment, for consistent policyholders whose working expenses are higher than rebellious contenders [4], [5]. To make solid business strategic, IMs seek to understand customer behaviors [6]. Therefore, this study attempts to (1) enquire the domain requirements for understanding customer claim behavior sequences, (2) understand whether customer claim behavior influence in the life insurance, and (3) investigate the importance of the precise formal and informal information causes from which individuals get financial knowledge.

Computer aided many problems-solving techniques have been extensively applied to find individual behaviors, such as dynamic adverse selection, account manipulation and false invoices [5], [7]. These approaches focus on detecting individual adverse records. However, they cannot analyze customer sequential claim behaviors, where many customers are involved in claiming their benefits [8]. Thus, the effective sequential mining and querying of individual records are of significant role in reducing the claim risk.

However, it is quite complex and challenging to understand, explore, and inspect the individual's potential claim behavioral issues [9], [10]. First, the identification of claim behavioral pattern, which is depends on the relationship between different customer and their various claim-related attributes. Second, the uncertainty in the claim history has arisen in the review procedure. Third, correlating an extensive amount of claims records and their analysis is time-consuming.

To handle the above challenges, we work with the insurance collaborator in Australia. Our collaborators provided policyholder claim records used in this paper to understand and explore the detailed requirements of analysts. We examined the related claim data, comprising policy holder profile, state, suburb, claim date, and claim-cause. We combine domain expertise with computer computing capacity and the expressiveness of visual analytics. We present *UCBVis*, an interweaving pattern mining and querying approach as shown in Figure 1, to help IMs understand, explore, and inspect the potential claim issues of customer. We then apply *UCBVis* to find possible activity that was raised in response to our collaborators' requirements. After that, we used this visual design to aid IMs in analysing customer behaviour. In addition, we used a user analysis and expert interviews to show the efficacy of this method on a real-world assertion dataset. We can summarize our contributions:

- We enquire the domain requirements for analyzing customer behavior, together with five domain expert's feed-

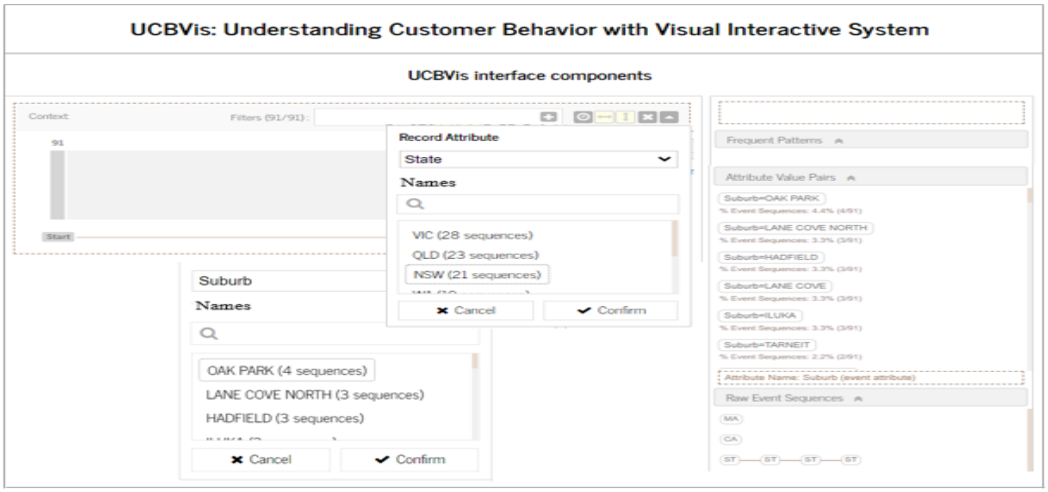

Fig. 1. Summary of *UCBVis* interface components. (a) Behavior exploration workflow: allow user to identify potential behavior of customer, (b) Frequent pattern view: allow user to choose sequential patterns for the focus, (c) Attribute pair view: allow user to search attribute names with behavioral patterns, and (d) Raw sequences view: can be supported with visualization.

back.

- By interweaving pattern mining and querying with an interactive visualization, we explore and inspect behavioral sequences from the life insurance claim records.
- We examine tasks of business analysts who aim to understand customer diverse behavior and visualize the outcomes.
- We report a user study with a large dataset and measured professional statements, which show the strength and usefulness of *UCBVis*.

We organized the paper as follows: Section II discusses related work on customer behavior exploration, techniques, and behavioral data visualization followed by the detail methodological discussion in Section III. The description of the visual analytic solution (*UCBVis*) is presented in Section IV. In Section V, we illustrate user study of the proposed visual analytics system (VAS) to assess and discuss its capacity to inform the relevant variables for exploring customer behavior in Section VI. Finally, conclusions and future directions are provided in Section VII.

II. Related Work

In this section, we review existing behavioural problems in the insurance industry as well as recent strategies most relevant to our work, such as consumer behaviour detection and data visualisation techniques for insurance claims.

A. Understanding customer behavior

The understanding of customer behavior is crucial because it has a high demand to the authorities for business management, product marketing, and decision making [11]–[13]. For instances, according to Islam et al. [5], by understanding and visualizing policyholders' claims data, IMs can avoid frauds and to provide risk management in the life insurance industry. Decision makers learn about inter-business activities and can develop new strategies [14]. To monitor a specific stock market user behavior, [15], [16] has provided adverse trading patterns and to identify the real-time stock marker performances. Additionally, to detect user adverse behavior, the coordinated specific keywords visualization is developed within the wire transactions [17]. However, there are very few works on exploring customer claim behavior sequences in the insurance domain. The existing systems have limitation to investigate many variables and satisfy specific requirements, e.g., measuring new claims costs, number of accidental claims, and number of mental health claims of domain experts.

B. Behavior analysis techniques

The majority of current literature on customer behaviour analysis is divided into two categories: 1) machine learning and 2) pattern mining techniques.

Machine learning methods: These approaches are based on developing models to measure accuracy for a variety of purposes [18]–[20]. For example, Amin et al. [21] has claimed that neural networks, decision trees, and Bayesian networks would help in successful evaluation of the impact of suspicious behavior in the business industry. From the findings of their study, it has been determined that suspicious scores of taxpayers are calculated from the information of taxpayers and their related invoice data. Ahmad et al. [22] depict the significance of real-world data for use of modelling as it accounts for more realistic cases. We can see from their research that they propose a hybrid model that combines the support vector machine, multi-layer perception neural network, and logistic regression classification models to detect tax evasion. Through the Wassouf et al. [23] study, we discovered that a multi-class predictive tool was created to detect fraudulent financial misstatements. Hglund [24] investigates the most influential features in predicting fraudulent tax payment behavior. They state that decision support tool is crucial in reducing fraud tax payment defaults. However, they cannot visualize the identification of insurance claim behavior records.

Pattern mining techniques: Pattern mining aid in discovering interesting insights by utilizing various data sequences [25]–[27]. For example, Islam et al. [5] propose association rule mining technique to represent the customer adverse behavior. They summarized adverse patterns and presented an unexpected behavioral records to explore suspicious user in the life insurance. Apaolaza, Aitor and Vigo, Markel [28] proposed a hybrid method for identifying suspicious behaviors that using pattern mining model impact performance and that clearance of identifying tax-burden can be evaluated. Graph based components are generally considered having differences between the diversity and accuracy. Thus, Ordonez et al. [29] investigate an attributed-graph based approach for large scale of claim data is used to detect suspicious records with rule mining technique. Although existing work on combining mining with querying is not a new idea, there has been no systematic investigation into customer behavior exploration in the insurance domain. Thus, our work lies at the intersection of both lines of research.

C. Behavioral data visualization

To provide possible insights into financial sector applications, we have categorised related financial data visualisation works into three categories: 1) time-series, 2) multi-attribute, and 3) financial fraud data visualisation

1. Time-series data: To detect, explore, and predict the many specific problems, there have been developed many visual analytic solutions for considering time-series data in diverse fields [30]–[32]. For example, a novel visualization system is developed for analyzing spatio-temporal correlations by Malik et al. [33]. Xie et al. [34] introduced a visualisation framework called VAET to distinguish salient transactions from large e-transaction time-series. Yue et al. [35] presented a novel timeline visualisation to look at the evolution of Bitcoin transaction trends from two viewpoints. In addition, existing work on visualising time-series data aims to investigate the interrelationships between them [36], [37]. However, there are some limitations embedded within existing research are required of data visualization for exploration and analysis.

2. Multivariate data: Multivariate data is employed for the spontaneous and interactive topological data inquiry [38]. To promote global analysis with strong communications to reinforce collections and aggregations, a large-scale multivariate network visualization system is introduced by Dai et al. [39]. Soriano et al. [17] presented simple visual encoding systems to concentrate on exploring the local subgraphs. In another study, they focus on investigating any unusual phenomenon to discover the most suspicious links and to simplify its identification.

3. Financial fraud data: The visual analytic system (VAS) has showed its relevance in identifying financial fraud [40], [41]. Existing work by Huang et al. [15] used visualization technique to detect fraud in the financial market. In addition, Qu et al. [42] proposed a VA framework to observe real-time stock market activity, control a specific stock, and use 3D treemaps to produce unusual pattern. In another study, Singh, Kishore and Best [43] demonstrated VISFAN, a visual analytics framework for detecting financial crimes like money laundering and fraud in financial activity networks. EVA is a set procedure with interactive visual analytics facilities proposed by Leite et al. [41], which offers a set procedure with interactive visual analytics facilities to play fraud confirmation.

In summary, our research is unique in that it focuses on customer behavior sequences. Thus, to discover customer claim behaviour problems, we presented multiple organised visualisations with carefully developed visual encoding schemes.

TABLE I
DEMOGRAPHIC CHARACTERISTICS OF CLAIM DATA SET

Variables	Category	Number of policyholders	Percentage (%)
Policy ID	Number	13287	100%
Claim-cause	Depression	326	2.45%
	Stroke	316	2.37%
	Neurotic disorder	1891	14.23%
	Accidental Falls	6192	46.60%
	Motor Vehicle Traffic Accidents	777	5.85%
	Cancer	2685	20.2%
	Unemployment	1364	10.26%
State	NSW	3566	26.83%
	VIC	2918	21.96%
	QLD	4099	30.85%
	SA	999	7.52%
	WA	1394	10.49%
	TAS	311	2.34%
Suburb	Whole Australia		
Date	2010-2019		

III. Methodology

This section introduces the design process by describing the data, domain goals, and requirements that experts hope to accomplish with the visualisation framework.

A. Data description

In this study, we collected a large scale dataset from one of the local Australian life insurance companies to study customer behaviour analysis. The dataset consists of claims record in over ten years periods (2010 to 2019) from 13287 countrywide policyholders. Each record comprises four attributes: ID (individuals unique ID), Claim-cause (what is the claiming reasons), State/Suburb (where they claim is recorded), Date (when and how frequently they are claimed). The detail of the dataset is described in Table I, in total, we have 13287 records.

B. Domain goals and requirement analysis

To better understand the customer behaviour and explore insights, we contacted with domain experts, who have been working over the past five years in the insurance industry. The experts have solid experience, and they are proficient for claim risk management in Australia. Face-to-face interviews and private meetings with domain experts were also held. We are interested in three aspects of customer behaviour analysis:

- What is the standard practice and procedure for identifying and analyzing customer behaviours?.

TABLE II
DOMAIN SPECIFIC REQUIREMENTS

SN	Requirement
R1	The system should allow analyses to be performed over selected suburbs of the city (the whole city or a part of it, e.g., Redfern in Sydney).
R2	The system should allow the types of claim-cause to be analyzed individually.
R3	The system should be simple and intuitive, allowing any officer to operate it with minimal training and basic computer skills.
R4	The system should allow for annotations and sharing, also allowing for continuous and shareable analyses.

- What are the main challenges and limitations of the ongoing approaches for detecting and analyzing customer claim behaviours?.
- What kind of study concerns and tasks do they prefer to bring?.

Through the interview sessions, the experts analyzed the current analysis strategy, which comprises three steps:

- Customers' period-end claim positions.
- Require new fieldwork and proper business records.
- Analysis of insightful information requires a vast amount of manual controlling of the raw claim data, which is tedious and time-consuming.

The above design requirements are shown in Table II.

C. Design Process

The paper is not about designing a new visual analytics system (VAS), rather considering an existing interactive VAS [44], we play a fundamental role in supporting the analysis task of business domains which requires thorough task analysis and domain expertise. We customize the design framework named *UCBVis* of an existing system for analyzing customer behaviour and to help the insurance risk manager. Thus, by following the study requirements discussed in Table II, *UCBVis* first identifies customer behavior sequences and then visualizes the outcome. Our system integrates pattern mining and querying techniques with an interactive visual interface. As shown in Figure 2, *UCBVis* is made up of three main modules: 1) data preprocessing, 2) data analysis, and 3) Visual Analysis.

The Data processing module performs both data masking and model construction. As mentioned in Subsection III-A, we record various information in the dataset with comprising various attributes. Since attributes have values in different categories, it may contain missing values. To simplify the system to ensure we use only the most significant data, data preprocessing comprised reducing less important and redundant attributes that offer no benefit to exploration and analysis. As part of data preprocessing, redundant fields that were not eliminated were combined. The Data Analysis module interweaves the mining and querying based on the parameters used to find novel insights from the insurance data. We allow analysts to evaluate behaviour sequences based on the exploration goals. To make it easier for people to explore themselves, we extract potential patterns from the data, and these patterns characterize individual behaviours of claim issues. Via intuitive visualisations, the visual analysis module introduces customers' different behavioural patterns, as well as the underlying contexts and comprehensive proof. It simplifies an accessible analysis and in-depth search of individuals with these related visualization views. We have presented these modules as a web-based system.

D. Implementation

The *UCBVis* system is developed as a web-based application, where Python is used to develop the backend to support data processing and analysis. JavaScript is used to implement the frontend where Data-Driven Documents (D3) is used to build visualization views. The interface is made up of HTML/Scalable Vector Graphics (SVG) components, and the web application is structured using the AngularJS framework, which follows the model-view-controller paradigm.

IV. Description of *UCBVis*

As shown in Figure 2, *UCBVis* comprises four major views: (a) Behavior extraction workplace, (b) the frequent pattern view, (c) the attribute-value pair view, and (d) raw sequence view. In this section, we describe how each workplace supports the user analytic tasks.

A. Behavior exploration workplace

The "Behavior exploration workflow" as shown in Figure 2 helps analysts who do not have an exact idea about how a VAS is being used to showing the full dataset as the context. It comprises several functionalities such as top bar, flow visualization, timeline. In order to promote the exploratory analysis, analysts can select "state" information, choose various "suburb area" that remain in the dataset to the flow visualization.

B. Incorporating frequent pattern into the workflows

As illustrated in Figure 2, the "Frequent Pattern View" component provides diverse information that is recorded various attributes and applied as input for the pattern mining algorithms. There are two different forms to get behavioral information from the frequent pattern workflow. First, users can use "Query Inputs" of the executed queries from the interface to take as input the results. Second, the system obtains a set of attribute under the corresponding header in the pattern analysis component, which includes the claim-cause and users' other information discussed earlier. Once the corresponding inputs are selected, users can get the raw interaction data through the "Attribute name selection" functionality. Additionally, user can set the minimum support parameters.

C. The attribute pair view

The extracted values are shown as a list in descending order of the level of sequences in the centre that contain claim records in the "attribute pair view." The list is made up of attribute pairs that all belong to the same attribute. By tapping on the attribute name, analysts may change the characteristic. When analysts hover over a claim attribute pair, a button

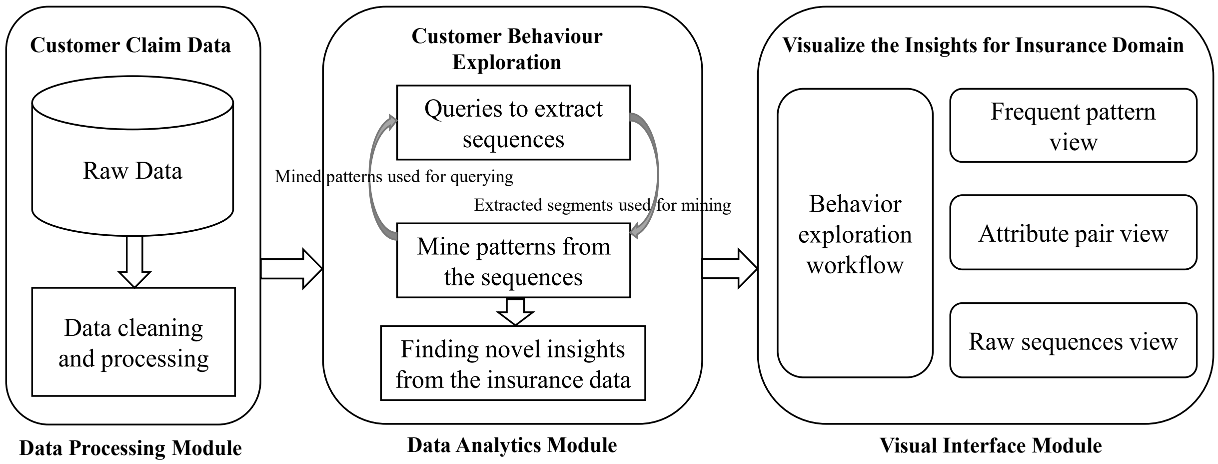

Fig. 2. The architecture of *UCBVis* system.

appears, allowing them to break the emphasis by that attribute pair. In addition, when you click this button, a new panel will appear with the focus separated and visualised.

D. Raw sequence view

In the "Raw sequence view", analysts can extend the chances of locating relevant patterns by customising the behavior set to be applied as input for pattern mining. We can utilize the behavior selection to adjust the input, putting on specific sequences of behaviors. Additionally, custom behaviors set up using the "Query Search" view by any user of the system, can also be added in the behavior set.

E. Interactions

Several core functionalities are combined in the *UCBVis* scheme. It supports focus+context, selection through dynamic filters and views, user-driven summaries, and multi-selection. Moreover, it applies animations to facilitate transition from temporal to sequential previews and join all attributes of information within the system, easing brushing and linking. By taking them above or below, the *UCBVis* structure can be further structured. The frequent pattern view can be applied to all entries in the research pane to further filter the claim records. Filters are cumulative and carry on to subsequent tabs. This system was used to assist users in drilling down and examining data. We also offer multi-selections, which enable users to select as many options as they want. We also provide multi-selections as it supports users to choose as many filters within and across tabs.

V. Evaluation

Participant: To observe how the visual responses generated, we conducted a user study with five domain experts' opinions which helps us to understand the penitential usefulness and utility of our system. All participants have been 25+ years and had executive experience in diverse fields. They were familiar with the importance of customer behaviour in the business industry. Although nobody offered any compensation and tried to use the system earlier, we ensured them to access unlimited afterwards.

Task and procedures: Detail study was conducted to assess how to visualize customer behavior to gain a better insights. In order to investigate the potential usability and effectiveness of *UCBVis*, we asked participants to respond to some open-ended questions about the system are given in Table III. We followed this up with additional feedback about the potential and desired capabilities. Afterwards, we suggested participants to run the system and give comment about the convenience and usability. A qualitative user study is preferred, integrated by a quick survey, as we believed details insights would be strongest through in combination with a fair interview. We explored to determine possible usability and accessibility issues, examine what views of the system would, and they would not use, decide what facilities required the most value, and describe differences and missing capabilities.

Results: In Table III, we selected the provided key observation questions to contain satisfactory statements. We present the outcome of the user study showed both promising results and constructive observation of our system. The users' comprehensive judgment was that the system would effectively understand, investigate, and measure customer behaviours. The exploration of an extensive database for the behavior analysis can provide useful insights for various purposes. Afterwards, we collected free form responses about what the participants considered relevant to our system's usefulness. In Figure 3, we show a variety of questions and ratings, where different user are provided different ratings. We observe that *UCBVis* uses can help the analysts gain attention for specific risks and provide the way how to deal with these risks. However, several impressive opinions were raised, such as the capability to store an image of a portfolio for a statement, the strength to annotate a portfolio. While these functionalities were not set up in our system, we will consider them for future versions.

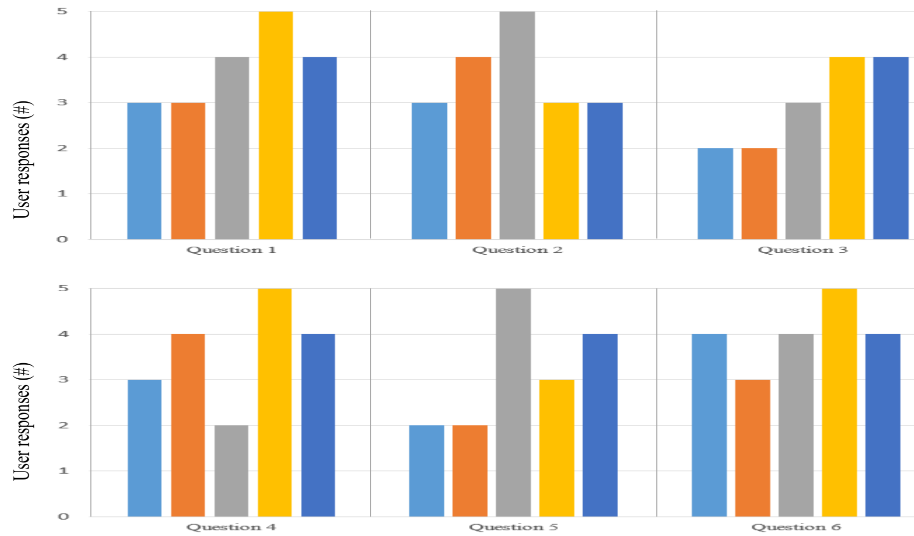

Fig. 3. User responses for the user study questions

TABLE III
USER STUDY RESULTS

Number	Category	Question	Mean (μ)	Std. Dev (σ)	Min.	Max.
Q1	Easy to use	UCBVis was easy to learn and use.	3.8	0.75	3	5
Q2	Insight	UCBVis was useful to explore insights patterns.	3.6	0.80	3	5
Q3	Insight	UCBVis allowed me to discover insightful queries about the data.	3	0.89	2	4
Q4	Essence	UCBVis helped me to generate knowledge about the claim data.	3.6	1.02	2	5
Q5	Speed	UCBVis enabled me to find interesting insights from the data quickly.	3.2	1.16	2	5
Q6	Confidence	UCBVis helped me to grow confidence about the interesting data insights.	4	0.63	3	5

VI. DISCUSSION

In Figure 3, we present the results of various questionnaires, where participants rated on a standard five points. We noticed that most of the responses were given by 3 and 4 points. The majority of the participants allowed that the system is handy for exploring visualizations and it permitted them to discover interesting insights from the data. As the user study and expert interviews have confirmed the capability and versatality of *UCBVis*, we provide a set of insights that can be helpful when deciding. In summary, we deduce the following key insights and guidelines:

- *UCBVis* can be used to understand customer diverse behavior sequences and inspect how past records affected in the industry.
- *UCBVis* allows IMs to search through the data by Suburb, State for exploring and inspecting behavioral pattern from the life insurance claim records.
- The primary risk information should be distinguished from secondary or less important information.
- Use of unnecessary elements in a visualization may cause confusion because of various expectations.
- Our user study report enabled to grow in confidence about the effectiveness and usefulness of *UCBVis*.

While the above mentioned insights ensure the effectiveness and use-fullness of *UCBVis*, it has some limitations, and some design choices still need further clarification. Therefore, we notice the following three key limitations that will help to confirm the effectiveness of *UCBVis* as follows:

1. Scalability: Scalability is an important aspect that we considered during the development of *UCBVis*. Due to the limited claim records, *UCBVis* works well for customer behavior analysis. However, when there are significantly more claim records, it may difficult to show the result in visual clutter overview. In terms of visualization, an improvement could be displayed.

2. Data availability and design choice. The system relies on the expert feedbacks where they stated that the policy premium of customer plays a significant role in claim inspection. However, in this study, the policy premium is not available in the dataset. Once the policy premium information is accessible, the system can provide more concrete evidence.

3. Visualization designs and usability: The visual interface of *UCBVis* is simple and easy to understand. For instance, the overview of the behavioral patterns employs the visual design of claim records. These visualization designs are straightforward. However, a VAS with better measured intuitive forms will be very simpler to find out the usability of *UCBVis*.

VII. Conclusion and Future Work

In this paper, we presented *UCBVis*, an interweaving pattern mining with querying through an interactive visualization system to help IMs for understanding customer claim behavior sequences. We described the method how the claim records can be understood by visualizing behavior sequences in a system. We described the strength and usability of *UCBVis* through a user study using real-world claim dataset.

In future work, we will deal with having large dataset that will include the post period of COVID-19 data. Also, it is possible to identify suspicious users, which domain experts would like to investigate. By integrating the used advanced sequential pattern matching tool, the visual description of our framework can be expanded. Last, we would like to test the usefulness of *UCBVis* with an eye tracking model, which builds an understanding of human perception, cognition, and interaction in order to design environments.

Acknowledgment

This work is partially supported by the Australian Research Council (ARC) under Grant No. DP200101374 and LP170100891. We would like to thank the Australian insurance company for providing us with the unique dataset and the authors of MAQUI system for supporting us to run the code.

References

[1] F. Taghikhah, A. Voinov, N. Shukla, and T. Filatova, "Exploring consumer behavior and policy options in organic food adoption: Insights from the australian wine sector," *Environmental Science & Policy*, vol. 109, pp. 116–124, 2020.

[2] S. Koldijk, J. Bernard, T. Ruppert, J. Kohlhammer, M. Neerincx, and W. Kraaij, "Visual analytics of work behavior data-insights on individual differences," 2015.

[3] C. Lin, Y.-J. Hsiao, and C.-Y. Yeh, "Financial literacy, financial advisors, and information sources on demand for life insurance," *Pacific-Basin Finance Journal*, vol. 43, pp. 218–237, 2017.

[4] S.-F. Chou, J.-S. Horng, C.-H. S. Liu, and J.-Y. Lin, "Identifying the critical factors of customer behavior: An integration perspective of marketing strategy and components of attitudes," *Journal of Retailing and Consumer Services*, vol. 55, p. 102113, 2020.

[5] M. R. Islam, S. Liu, R. Biddle, I. Razzak, X. Wang, P. Tilocca, and G. Xu, "Discovering dynamic adverse behavior of policyholders in the life insurance industry," *Technological Forecasting and Social Change*, vol. 163, p. 120486, 2020.

[6] W. Didimo, L. Grilli, G. Liotta, L. Menconi, F. Montecchiani, and D. Pagliuca, "Combining network visualization and data mining for tax risk assessment," *IEEE Access*, vol. 8, pp. 16 073–16 086, 2020.

[7] M. R. Islam, I. Razzak, X. Wang, P. Tilocca, and G. Xu, "Natural language interactions aided with data visualization for exploring insurance claims and risk management," *Annals of Operations Research*, 2021.

[8] S. K. Roy, R. L. Gruner, and J. Guo, "Exploring customer experience, commitment, and engagement behaviours," *Journal of Strategic Marketing*, pp. 1–24, 2020.

[9] T. Lang and M. Rettenmeier, "Understanding consumer behavior with recurrent neural networks," in *Workshop on Machine Learning Methods for Recommender Systems*, 2017.

[10] A. Vij, S. Ryan, S. Sampson, and S. Harris, "Consumer preferences for on-demand transport in australia," *Transportation Research Part A: Policy and Practice*, vol. 132, pp. 823–839, 2020.

[11] A. Yaeli, P. Bak, G. Feigenblat, S. Nadler, H. Roitman, G. Saadoun, H. J. Ship, D. Cohen, O. Fuchs, S. Ofek-Koifman *et al.*, "Understanding customer behavior using indoor location analysis and visualization," *IBM Journal of Research and Development*, vol. 58, no. 5/6, pp. 3–1, 2014.

[12] S. Lee, C. Min, C. Yoo, and J. Song, "Understanding customer malling behavior in an urban shopping mall using smartphones," in *Proceedings of the 2013 ACM conference on Pervasive and ubiquitous computing adjunct publication*, 2013, pp. 901–910.

[13] F. Tian, T. Lan, K.-M. Chao, N. Godwin, Q. Zheng, N. Shah, and F. Zhang, "Mining suspicious tax evasion groups in big data," *IEEE Transactions on Knowledge and Data Engineering*, vol. 28, no. 10, pp. 2651–2664, 2016.

[14] F. Du, C. Plaisant, N. Spring, and B. Shneiderman, "Finding similar people to guide life choices: Challenge, design, and evaluation," in *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, 2017, pp. 5498–5544.

[15] M. L. Huang, J. Liang, and Q. V. Nguyen, "A visualization approach for frauds detection in financial market," in *2009 13th International Conference Information Visualisation*. IEEE, 2009, pp. 197–202.

[16] T. Fujiwara, N. Sakamoto, J. Nonaka, K. Yamamoto, K.-L. Ma *et al.*, "A visual analytics framework for reviewing multivariate time-series data with dimensionality reduction," *IEEE Transactions on Visualization and Computer Graphics*, 2020.

[17] A. Soriano-Vargas, B. Hamann, and M. C. F de Oliveira, "Tv-mv analytics: A visual analytics framework to explore time-varying multivariate data," *Information Visualization*, vol. 19, no. 1, pp. 3–23, 2020.

[18] M. R. Islam, S. Liu, X. Wang, and G. Xu, "Deep learning for misinformation detection on online social networks: a survey and new perspectives," *Social Network Analysis and Mining*, vol. 10, no. 1, pp. 1–20, 2020.

[19] M. R. Islam, M. A. Kabir, A. Ahmed, A. R. M. Kamal, H. Wang, and A. Ulhaq, "Depression detection from social network data using machine learning techniques," *Health information science and systems*, vol. 6, no. 1, pp. 1–12, 2018.

[20] M. R. Islam, A. R. M. Kamal, N. Sultana, R. Islam, M. A. Moni *et al.*, "Detecting depression using k-nearest neighbors (knn) classification technique," in *2018 International Conference on Computer, Communication, Chemical, Material and Electronic Engineering (IC4ME2)*. IEEE, 2018, pp. 1–4.

[21] A. Amin, F. Al-Obeidat, B. Shah, A. Adnan, J. Loo, and S. Anwar, "Customer churn prediction in telecommunication industry using data certainty," *Journal of Business Research*, vol. 94, pp. 290–301, 2019.

[22] A. K. Ahmad, A. Jafar, and K. Aljoumaa, "Customer churn prediction in telecom using machine learning in big data platform," *Journal of Big Data*, vol. 6, no. 1, p. 28, 2019.

[23] W. N. Wassouf, R. Alkhatib, K. Salloum, and S. Balloul, "Predictive analytics using big data for increased customer loyalty: Syriatel telecom company case study," *Journal of Big Data*, vol. 7, pp. 1–24, 2020.

[24] H. Höglund, "Tax payment default prediction using genetic algorithm-based variable selection," *Expert Systems with Applications*, vol. 88, pp. 368–375, 2017.

[25] T. E. Meyer, M. Monroe, C. Plaisant, R. Lan, K. Wongsuphasawat, T. S. Coster, S. Gold, J. Millstein, and B. Shneiderman, "Visualizing patterns of drug prescriptions with eventflow: A pilot study of asthma medications in the military health system," OFFICE OF THE SURGEON GENERAL (ARMY) FALLS CHURCH VA, Tech. Rep., 2013.

[26] Y. Chen, Z. Zheng, S. Chen, L. Sun, and D. Chen, "Mining customer preference in physical stores from interaction behavior," *IEEE Access*, vol. 5, pp. 17 436–17 449, 2017.

[27] Z. Liu, Y. Wang, M. Dontcheva, M. Hoffman, S. Walker, and A. Wilson, "Patterns and sequences: Interactive exploration of clickstreams to understand common visitor paths," *IEEE Transactions on Visualization and Computer Graphics*, vol. 23, no. 1, pp. 321–330, 2016.

[28] A. Apaolaza and M. Vigo, "Assisted pattern mining for discovering interactive behaviours on the web," *International Journal of Human-Computer Studies*, vol. 130, pp. 196–208, 2019.

[29] L. Ordonez-Ante, G. Van Seghbroeck, T. Wauters, B. Volckaert, and F. De Turck, "Explora: Interactive querying of multidimensional data in the context of smart cities," *Sensors*, vol. 20, no. 9, p. 2737, 2020.

[30] K. Li, Y.-N. Li, H. Yin, Y. Hu, P. Ye, and C. Wang, "Visual analysis of retailing store location selection," *Journal of Visualization*, vol. 23, no. 6, pp. 1071–1086, 2020.

[31] Y. Du, C. Wang, C. Li, and H. Yin, "Behaviortracker: Visual analytics of customer switching behavior in o2o market," in *Proceedings of the 11th International Symposium on Visual Information Communication and Interaction*, 2018, pp. 17–24.

[32] K. Takai and K. Yada, "A framework for analysis of the effect of time on shopping behavior," *Journal of Intelligent Information Systems*, vol. 41, no. 1, pp. 91–107, 2013.

[33] A. Malik, R. Maciejewski, N. Elmqvist, Y. Jang, D. S. Ebert, and W. Huang, "A correlative analysis process in a visual analytics environment," in *2012 IEEE Conference on Visual Analytics Science and Technology (VAST)*. IEEE, 2012, pp. 33–42.

[34] C. Xie, W. Chen, X. Huang, Y. Hu, S. Barlowe, and J. Yang, "Vaet: A visual analytics approach for e-transactions time-series," *IEEE transactions on visualization and computer graphics*, vol. 20, no. 12, pp. 1743–1752, 2014.

[35] X. Yue, X. Shu, X. Zhu, X. Du, Z. Yu, D. Papadopoulos, and S. Liu, "Bitextract: Interactive visualization for extracting bitcoin exchange intelligence," *IEEE transactions on visualization and computer graphics*, vol. 25, no. 1, pp. 162–171, 2018.

[36] B. C. Kwon, J. Verma, and A. Perer, "Peekquence: Visual analytics for event sequence data," in *ACM SIGKDD 2016 Workshop on Interactive Data Exploration and Analytics*, vol. 1, 2016.

[37] P. C. Yau, D. Wong, W. H. Luen, and J. Leung, "Understanding consumer behavior by big data visualization in the smart space laboratory," in *Proceedings of the 2020 5th International Conference on Big Data and Computing*, 2020, pp. 13–17.

[38] M. T. Islam, M. R. Islam, S. Akter, and M. Kawser, "Designing dashboard for exploring tourist hotspots in bangladesh," in *The 23th International Conference on Computer and Information Technology (ICCIT-2020*. IEEE, 2020.

[39] W. Dai and M. G. Genton, "Multivariate functional data visualization and outlier detection," *Journal of Computational and Graphical Statistics*, vol. 27, no. 4, pp. 923–934, 2018.

[40] R. A. Leite, T. Gschwandtner, S. Miksch, E. Gstrein, and J. Kuntner, "Neva: Visual analytics to identify fraudulent networks," in *Computer Graphics Forum*, vol. 39, no. 6. Wiley Online Library, 2020, pp. 344–359.

[41] R. A. Leite, T. Gschwandtner, S. Miksch, S. Kriglstein, M. Pohl, E. Gstrein, and J. Kuntner, "Eva: Visual analytics to identify fraudulent events," *IEEE transactions on visualization and computer graphics*, vol. 24, no. 1, pp. 330–339, 2017.

[42] Z. Qu, C. W. Lau, D. R. Catchpoole, S. Simoff, and Q. V. Nguyen, "Intelligent and immersive visual analytics of health data," in *Advanced Computational Intelligence in Healthcare-7*. Springer, 2020, pp. 29–44.

[43] K. Singh and P. Best, "Anti-money laundering: Using data visualization to identify suspicious activity," *International Journal of Accounting Information Systems*, vol. 34, p. 100418, 2019.

[44] P.-M. Law, Z. Liu, S. Malik, and R. C. Basole, "Maqui: Interweaving queries and pattern mining for recursive event sequence exploration," *IEEE transactions on visualization and computer graphics*, vol. 25, no. 1, pp. 396–406, 2018.