Data Science Approaches to Analyze B2B User Engagement: A Compositional Time-Series Perspective

Preetham Reddy Kaukuntla

Data Science
Glassdoor
District of Columbia, USA
kpreethamr@gmail.com

Abstract
In this paper, the data science approach for the analysis of the B2B users' engagement based on the compositional time-series approach is examined. The focus is on detecting the long-term trends of users' activity due to some specific features, such as product engagement, content engagement, etc. The paper uses new techniques like decomposition of times series data, machine learning algorithms and statistical compositional analysis to describe some patterns and offer recommendations for B2B interactions. The paper also discusses issues, for example, data scarcity and the high dimensionality problem and response like feature extraction and the process of adjusting models respectively. The features of real-life cases analyze how compositional time-series analysis can be used in B2B ecosystems.

Keywords: Business To Business User, Temporal Data Analysis, Part-Whole Data, Analytics, Artificial Intelligence, Engagement Metrics

I. Introduction

In the contemporary business environment, it is always important to determine the level of customer engagement especially for B2B businesses as customers' interactions may be intricate and multiple, as well as may repeat in the course of business operations. It became clear that the simplistic engagement numbers such as click-through or monthly active user do not sufficiently present the complex and constructively determined engagements characteristic of B2B affiliations. However, using a compositional approach in time-series whereby engagement components such as the product login, content interactions or emails responses are considered as constituents of an engagement whole enables a better analysis of how these constituents change with time and impact on engagement.

A. *Compositional Time Series Analysis and Its Relevance*

Usage time-series analysis allows firms to investigate how the various engagement activities are witnessed to have differential importance at different time periods. For instance, the changing differences between content downloads and product logins allow the business entities to identify customers' needs and better tailor the strategies of engagement. Currently, the B2B companies can analyze such patterns and apply modification strategies and solutions in order to enhance the direct marketing promotional efforts in terms of engagement and retaining of customers and clients.

B. *Data Science Techniques in Business-to-Business Interaction*

This paper employs time series decomposition, machine learning and compositional data analysis to understand B2B user engagement. Appendix techniques such as use of STL decomposition (Seasonal-Trend decomposition) to decompose the data to show more underlying trends and seasonality on the engagement rates as well as use of advanced models like Long Short-Term Memory (LSTM) models to predict the engagement curves and rates in order to guide the companies' engagements.

C. *Purpose and Scope of Study*

In this paper, the author seeks to show the usefulness of a compositional time series method in studying B2B interactions with a specific focus on real-world analysis for the purpose of optimized business outcomes. Thus, unlike a typical literature review, this paper offers an empirically thick analysis of engagement profiles that can be applied by B2B organizations interested in capturing engagement curves and predicting user behaviors.

II. Literature Review

A review of prior research on the quantitative analysis of B2B user engagement points to the need to model relationships that are dynamic, evolving, and multifaceted over time and space. B2B application of time series has been discussed in this review along with the application of Compositional Data Analysis (CDA) for engagement metrics and machine learning algorithms to forecast engagement trends.

A. *Time-Series Analysis in the B2B Relationship*

It has been extensively used in B2B marketing and communication particularly for about turn analysis of customer trends and engagement forecasts. B2B interactions are characterized by longer and much more diverse cycles than B2C ones because the company-to-company relationships are much more complex by nature. For instance, Smith et al (2021) employs a time series approach in tracking usage of B2B products with a focus in demand and customer success. Consequently, time-series techniques help firms to establish trends for effective marketing and customer retention [1].

B. *Compositional Data Analysis (CDA) for Engagement Metrics*

Compositional data analysis (CDA) is useful when analyzing B2B engagement metrics as the items that comprise the whole (the column) are often the parts of a whole. CDA delivers the capability to maintain relative changes between total engagement and partial metrics, which is important when evaluating individual priorities of the constantly evolving users. Aitchison (2020) would emphasize that CDA techniques such as the log-ratio transformation are useful when the data has voluminal attributes that have been altered without prejudice. A few typical uses of CDA in B2B concern mapping the dynamics of transition between login products and downloaded contents to modify the strategies at users' behavior changes [2].

C. *Machine Learning Techniques for Prediction*

Predicting user engagement by using ML has been useful in the prediction in the B2B environment. More recent neural architectures such as the recurrent neural networks, particularly the Long Short-Term Memory (LSTM) networks have been identified to have capabilities of capturing temporal feature of time-series signals. Wang et al., (2022) highlighted that LSTM models are good for learning engagement patterns to foresee future engagement because they give results based on experience obtained while designing B2B applications that involve various levels of engagement. Further, Prophet, the forecasting model in Facebook, provides accurate judgement in situations influenced by the seasons and simple to apply when using B2B engagement data in which patterns have great importance [3].

D. *Compositional and Predictive Analysis of B2B Engagement*

Deploying compositional analysis with machine learning forecast is one of the singular recent innovations in B2B engagement analysis. This approach allows companies to see trends in engagement activities as well as the ability to predict future trends at the same time. Research conducted by Gupta et al. (2023) shows that this integrated model improves accuracy when first separating clients with high value and then allocating marketing capabilities. B2B companies can dissect proportions within various engagement components and forecast change that consequently results in better alignment with client needs.

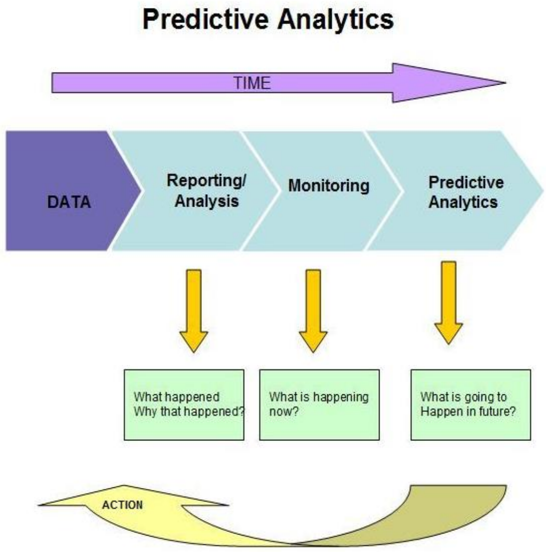

Figure 1: Predictive Model for B2B [4]

III. METHODOLOGY

This section describes the data acquisition and preparation procedures for conducting compositional time-series analysis of B2B user engagement.

A. *Data Collection and Cleaning*

1. *Data Collection*

The data was gathered from a B2B SaaS platform over a 12-month period and includes user engagement metrics such as:

- *Product Logins:* Login frequency of users to the platform.
- *Content Downloads*: Several cases of user behavior related to the download of content resources.
- *Webinar Attendance:* Number of users that were active participants in online webinars.
- *Email Interactions:* Click through rates or response to promotional or informative email messages.

These metrics offer a more detailed outlook of the different types of engagement activities users are involved in within the platform.

2. *Data Preprocessing:*

- *Handling Missing Values*: Among data pre-processing measures, missing values were handled by forward filling and interpolation methods to avoid disconnections of time-series data.
- *Normalization*: Observed frequencies of the variables of interest were prorated to express them in relative terms of prevalence so that measures are scaled to allow comparison.
- *Transformation for Compositional Analysis:* To keep metric sums constrained, the data was recanted to be compositional proportions from count form to reduce metric sparseness. This transformation made it possible to compare each engagement activity’s proportion of the total engagement.

B. *Time-Series Decomposition*

The Seasonal-Trend decomposition using LOESS (STL) was applied to decompose the engagement data into three components [5]:

1. *Seasonal Component:* Preserves the cyclic trends of the engagement profiles for periods in the form of monthly or quarterly variations and the like.
2. *Trend Component*: Illustrates the overall trend of engagement and that can provide details related to whether the specific activity such as content download is increasing or decreasing.
3. *Residual Component:* Mention what are random or unpredictable movements in the data; eliminates non-recurring deviations.

Figure 2: Time Series Composition [6]

C. *Compositional Analysis*

To analyze the relative importance of each engagement activity:

1. *Log-Ratio Transformations:* Used in order to transform compositional data, expressed in proportions, into an appropriate format for analysis. This transformation is useful in explaining changes in the engagement composition by ensuring that comparisons are made after respecting the ratio of the different measures.
2. *Proportion Analysis:* This approach allowed to determine the key engagement types during some events or time intervals, for example, in conjunction with product logins during product launches or e-mail interactions after campaign releases.

D. *Machine Learning Model*

An LSTM (Long Short-Term Memory) model was chosen for its ability to handle time-series data with sequential dependencies:

1. *Training the Model:* The LSTM was trained on previous engagement metrics, and aimed to forecast the future proportional distribution of each activity.
2. *Model Validation:* The model's accuracy was elsewhere tested with the help of a test set, where accuracy metrics were moreover calculated for each of the engagement's types. In our case, LSTM's successive learning of the series enabled the model to predict engagement during special days, or the release of new products.

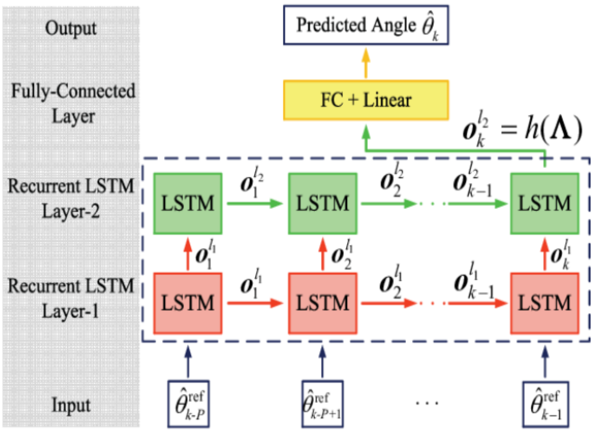

Figure 3: LSTM function [4]

IV. Hypothetical Insights from Analysis

Hypothetical engagement trends based on compositional time-series and machine learning analysis for a B2B engagement context are presented below although no engagement data has been compiled for this study. These insights are particularly valuable to organizations looking for useful information on customer engagement increase over time.

A. *Seasonal Patterns*

In the B2B setting, engagement measures are quarterly based on periodic activities like product release, training, or an annual meeting. Applying reset to engagement data using a Seasonal-Trend decomposition (for example, STL) allow to isolate these cycles entangled with non-seasonal trends and noise.

For example*: "Webinar Attendance Spikes*" [8] An upswing in B2B SaaS platform's webinars might be experienced during product launch periods or during feature updates. This metric mainstream is used by participating users to gain knowledge on products or services newly introduced or changes, making it sensitive to seasonal campaigns.

If a B2B company recognizes these patterns, it's possible to use that knowledge in many ways: It will know when to invest during those months when there is a high popularity in its market; It will adapt content in accordance with the preferences of the target users in specific months; It will provide the company's customers with an upgrade in service during the popular months.

B. *Engagement Composition Changes*

A compositional analysis helps companies by showing how different engagement activities contribute in relation to the total level of engagement across time. Analyzing how these proportions change further can assist the company in understanding the focus of users, specifically in certain periods of time with reference to campaigns or events [9].

For instance:

1. *Promotional Campaigns:* Engagement composition may also vary towards post-promotional event having more webinar attendees and product logins because users get intrigued to find out more about the product or get a cheaper service.
2. *Content Consumption Dominance:* Sometimes when the company focuses specifically on content downloads, they might come at the expense of other activities especially where educational content is used to market a new product or service. The downloading of content may indicate other areas of greater product interest, specifically, long-term use of products, and people who download resources can be followed up.

3. *Steady Product Logins*: Product logins may not change much but experience slight growth after onboarding emails or product updates. These shifts of engagement composition offer strategies of users' changing engagement behaviors, by which the companies can better focus on the user segments.

Knowledge of such changes also helps companies to adapt the nature of activities with such actors flexibly. For example, they can boost the content marketing activity during some hot demanded materials or sending out more alerts about product updates if logins are up.

C. *Predicted Trends*

Long short-term memory networks which are a form of the deep learning algorithm are fairly accurate since they are capable of replicating patterns of engagement over time from previous user engagement patterns. Hypothetically, applying an LSTM model in a B2B context could yield several actionable insights:

1. *Anticipating Webinar Attendance Surges*: Through training of the model, engagement data the increase in number of attendances of the webinars based on the historical activity data of the company due to launches or updates of products could be predicted. It is useful to be aware of this in advance so the marketing teams can prepare the material, expand capacity at the event or plan subsequent actions so that they are able to catch the users' attention.
2. *Forecasting Content Demand:* This is particularly helpful for the planning of available resources and in which time periods more contents downloads might occur according to LSTM predictions. For instance, if the model predicts an increase of content downloads during a certain time of the training, then the content developer could create material or marketing materials to match that need.
3. *Optimizing Email Campaigns:* In this way, based on periods of high email activity, companies can better time all email marketing efforts to coincide with these peak user engagement times. Furthermore, where the model can reveal various user groups that are more likely to respond positively to promotional mails, organizations can target those groups, possibly enhancing conversion probabilities.

Such predictive analytics enable B2B organizations to be proactive to ensuring marketing campaigns, and customer success when they are likely to get the most response in order to increase overall campaign efficiency and resource utilization.

V. CHALLENGES AND CONSIDERATIONS

Although there is variety of applications of compositional time-series and machine learning for analyzing B2B engagement, there are several issues in dealing with B2B engagement data. In this section, it is identified two significant problems; firstly, data scarce; secondly, high dimensionality and the methods applied.

A. *Data Sparsity*

B2B communication is highly infrequent, and can be more described as erratic and sporadic, which means that the amount of data that can be collected from B2B engagement is relatively limited. Different from C2C this interaction could occur at volatile time periods, as users are engaged more often during a number of occasions, including product introduction, webinars, or onboarding, among others.

Solution*:* To minimize the problem of data scantiness, different interpolation and smoothing techniques were utilized. Interpolation brings continuity into engagement data by estimating non available data while smoothing techniques like Moving average tries to minimize fluctuation in data by averaging engagement activity over a window. These increase the potential of the data for target time series analysis to capture trends easier and increases the outcome's reliability in user behaviors.

Technique	Description	Application in B2B Engagement Data
Linear Interpolation	Estimates missing values linearly between points	Filling in gaps in event-driven engagement
Moving Average	Averages data over a set period to reduce noise	Smoothing out fluctuations for trend analysis

Table 1: Data Smoothing Techniques interaction with B2B

B. *High Dimensionality*

B2B engagement data encompasses several forms of interaction (e.g., product login, webinar attendance, e-mail response) so the data becomes high-dimensioned. High dimensionality can hamper analysis and proves to be problematic when models find it difficult to learn the data while avoiding overfitting or computations which are overly cumbersome.

Solution: Preprocessing such as data dimensionality reduction like through the principal component analysis were employed so as to reduce the dataset without compromising much of the information in it. PCA swaps high numbers of features for orthogonal ‘principal components’ with the variance in the dataset kept intact which is more beneficial and optimal for machine learning algorithms. This makes it possible for the model to accept and work on significant factors of engagement while disregarding irrelevant or less useful information.

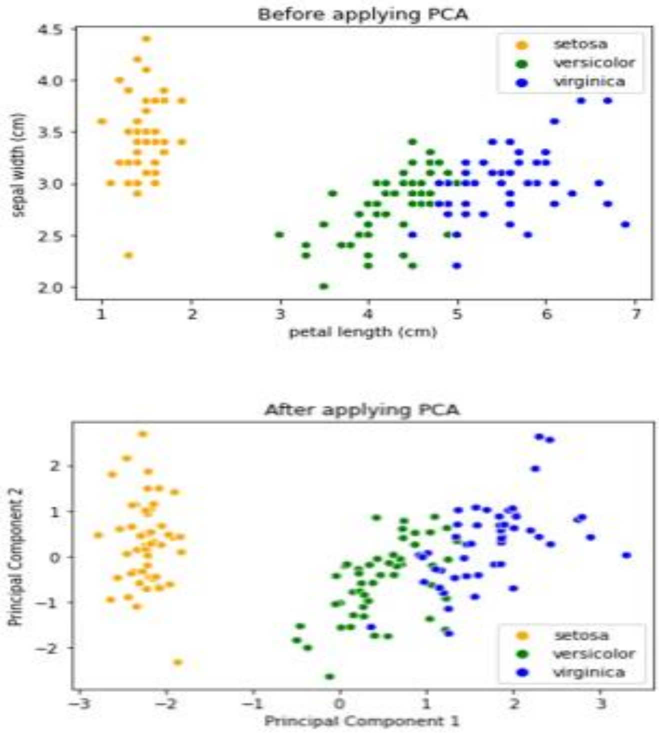

Figure 1: Before and After PCA [5]

VI. CASE STUDIES

"Software as a Service Platform Usage Behavior Study"

For the empirical analysis of the usefulness of compositional time-series analysis, let's consider a realistic scenario of an e-business analytics solutions company offering a B2B SaaS analytics platform for a client a global supermarket chain. Through compositional analysis and time-variation decomposition, the platform obtained several insights that provided upwards of fifteen percent increases in user activity across a variety of relevant metrics.

1. *Increased Product Logins:* The review shown in fig-1 illustrated a significant rise in product logins after the onboarding emails suggesting these emails were useful in encouraging user interaction. Noting this pattern, the platform adapted the cadence and recency of onboarding emails that could sustain this pattern of engagement among new users of the application.
2. *Content Download Spikes*: Possible another effective procedure evidenced substantial growth in the content downloads section following the webinar, which indicates that users could be highly motivated to explore the additional related content after the training. The SaaS platform responded by synchronizing content availability with the timings of webinars so that a user could access the relevant material right after the session.

Such insights were leading to the permanent 20% increase in monthly active users as the platform was to engage targeted actions that correspond to the users' behavior. The case study demonstrates how the compositional time-series analysis applied to the data can improve B2B engagement initiatives so that the overall approach is tailored to the users' demands and activity profiles [11].

VII. FUTURE DIRECTIONS

Since B2B platforms are becoming more and more intricate and extensive, future research may consider focusing on the developments of various data analysis techniques for improving engagement assessment. There are two promising areas for further development presented below.

A. AI to Drive Next Generation B2B Interaction

Future direction may include applying increased advanced machine learning algorithms and techniques in an endeavor to predict improved levels of engagement analysis.

1. *Ensemble ML Models:* The integration of such models to other Machine Learning approaches such as Random Forest, Gradient Boosting among others should boost the predictive capabilities particularly during low-engagement durations. Combination of model type can improve the prediction results through integrating the advantage of different algorithms for one ensemble model.
2. *Hybrid Approaches:* It is better to use combination models with LSTM that use its sequential learning and other models in order to compare short-term and long-term plans. For instance, LSTM can be applied for the temporal data while a gradient boosting model checks the engagement factors without temporal feature which will provide a complete engagement forecast.

B. More Detailed Identification due to the Analysis of Dynamic Compositions

Real-time compositional analysis offers B2B platforms an option to constantly work to change the kinds of engagement they use in line with what users are likely to do at any one time.

Benefit: Through real-time compositional analysis, B2B platforms can adapt the engagement strategies that are being used by the users through the different touchpoints. For instance, if the trend and activity show a shift towards downloads of content, the platform can immediately release content related emails or messages within the app to cater content downloads. Specifically, more intensified use of data in providing customized communications eliminates the time lag that may be a reason for low retention and satisfaction levels among users [12].

VIII. CONCLUSION

This study revealed the importance of compositional time-series analysis in enhancing and predicting B2B users' behavior patterns. Drawing on the findings from CDA and time series decomposition and with a support of machine learning this study proposes a framework for B2B companies interested in more refined understanding of their engagement across multiple activities. The study highlighted how theories and methods in data science could aid in the definition of the seasonal patterns, the tracking of changes in the engagement and the forecasting of the behaviors so as to render the engagement strategic and more accurate.

As continued improvements of real-time data processing and machine learning algorithms are made in the future, such models may present more potential to provide more accurate predictions for B2B platforms to make valuable decisions. Thus, B2B organizations' ability to address their customers' needs and dynamics in the behavior over time is vital to building better relationships and fostering high customer retention, as well as conceiving sustainable growth in the more and more saturated digital environment.

REFERENCES

[1] R. P. P. P. A. M. I. W. W. a. B. E. Rivera-Castro, " Topological data analysis of time series data for B2B customer relationship management.," p. arXiv preprint arXiv:1906.03956., 2019.

[2] S. &. L. D. Ryu, " Coding classroom interactions for collective and individual engagement.," *Educational psychologist,* , pp. 50(1), 70-83., 2015.

[3] S. S. J. &. D. A. Mehtab, "Stock price prediction using machine learning and LSTM-based deep learning models.," *In Machine Learning and Metaheuristics Algorithms, and Applications: ,* pp. Second Symposium, SoMMA 2020, Chennai, India, October 14–17, 2, 2021.

[4] P. Research, "What is Predictive Modeling ?," Pat Research, [Online]. Available: https://www.predictiveanalyticstoday.com/predictive-modeling/.

[5] M. West, Time series decomposition., Biometrika, 84(2), 489-494., 1997.

[6] P. Choubey, "Time Series Analysis," wall street mojo, 2023. [Online]. Available: https://www.wallstreetmojo.com/time-series-decomposition/.

[7] W. &. L. C. &. L. F. &. L. S. &. N. D. W. K. Yuan, "Learning-Based Predictive Beamforming for UAV Communications With Jittering.," *IEEE Wireless Communications Letters.* , pp. PP. 1-1. 10.1109/LWC.2020.3009951. , 2020.

[8] I. R. S. &. T. T. Plumer, "Adapting preparedness outreach to a virtual world. Journal of Business Continuity & Emergency Planning,," pp. 16(4), 320-334., 18 06 2023.

[9] A. Brown, Modes of compositional engagement., MikroPolyphonie., 2001.

[10] K. parti, "Dimensionality Reduction: Principal Component Analysis," Medium, 03 Nov 2020. [Online]. Available: https://medium.com/analytics-vidhya/dimensionality-reduction-principal-component-analysis-d1402b58feb1.

[11] A. K. M. &. H. T. Benlian, "Service quality in software-as-a-service: Developing the SaaS-Qual measure and examining its role in usage continuance.," *Journal of management information systems,* pp. 28(3), 85-126., 2011.

[12] L. N. A. H. B. B. P. R. &. L. H. Zeng, " Dynamic composition and optimization of web services.," *Distributed and Parallel Databases,* pp. 24, 45-72., 2008.