School of Informatics

Agent-based modelling for sustainable banking — collaboration with ekko Agent-based Modelling

B208489
May 2022

Abstract

Ecosystems are under threat from climate change. However, the link between reducing carbon emissions and spending behaviour exists, and the project aims to use the link to reduce the individual carbon footprint. The project uses an agent-based modelling approach to analyse people spending habits and discovers new influence maximisation strategies to encourage people to purchase at sustainable companies.

Date: Sunday 1st May, 2022

Tutor: Valentina Andries
Supervisor: Valerio Restocchi

1 Motivation

Nowadays, our attention to environmental issues has been increasing, and the environmental issues we face are getting more and more severe. One way that everyone can contribute to environmental protection in our daily life is to reduce their carbon footprint, which is the total amount of greenhouse gases caused by our actions [**?**]. When we buy different products in different places, greenhouse gas emissions also vary, but we can find a method to encourage people to buy products at sustainable companies [**?**]. Hence, we need to understand people spending habits and how to make valid influences. Agent-based modelling (ABM) is a simulation technique that captures individual behaviour in an environment that has been applied to many real-world problems in recent years, especially for business problems [**?**]. Most business problems are hard to simulate as we do not know the behaviour of the system or the system is too complex for conventional approaches. However, the ABM provides a way to simulate a large number of decisions by individual entities with simple decision rules [**?**]. If we successfully simulate the business problem, we could understand the inner logic of the system, how individual entities act and interact with each other and the market or even discover the reason that might cause business failure. Moreover, we can also change the rules of the ABM and observe the reflection of the model before we apply changes to the real-world problem to avoid loss. Therefore, in this project, we will focus on the spending habit of people. We aim to build an ABM that successfully simulates people's daily spending habits and find new influence maximisation strategies that affect people's spending habits based on the ABM. Then we could apply the influence strategies to encourage people to reduce their carbon footprint.

1.1 Feasibility

This is an industry-supported project, so the transaction data will be provided by ekko as spending habit data. We could use network science to analyse the behaviour of user groups and use the obtained information to build and test the ABM. Then, we could start to find influence maximisation strategies with our ABM.

1.2 Beneficiaries

With the actual data analysis and ABM model results, we could apply the strategies to the real world to encourage people to shop at greener companies. In addition, we could use our user spending habits results to help the industry provide more satisfying services for customers.

2 Background and Related Work

2.1 Network Science

Networks appear in all kinds of aspects of our life, so it becomes a powerful way to represent and analyse complex interactions [**?**]. A network can be treated as a graph where nodes can represent different individuals, and links represent interactions between them. According to the described system situation, the network can be classified into weighted or unweighted, and directed or undirected [**?**]. We could obtain a conceptional understanding of the system by analysing the network model, hub, community, and pattern.

2.2 Agent-based Modeling

The ABM provide a bottom-up way of describing the system and is a microscale model. In agent-based modelling, a system is built by autonomous agents. Each agent in this system has to make decisions based on a set of rules, and agents behave differently to represent different systems [**?**]. Therefore, the ABM aims to use multiply agents to reproduce or predict a complex phenomenon as the ABM simulates all agents' interactions simultaneously and also captures emergent phenomena [**?**]. The result of the ABM has high randomness and is hard to predict as we can only set the rules and agents' behaviour at the beginning, and we can not control the growth of the model. Hence, it is an excellent way to test our understanding of the system.

2.3 Related Work

Using an agent-based modelling approach to simulate a human involved system and analysis the human spending habits is not rare. For example, Andrea et al. provide an agent-based modelling simulation of customer behaviour in meat consumption in Britain with consideration of individual preferences and peer influence [**?**]. Agents are created with different gender, ages, salary, employment status, diet, and others. Each agent has a meat consumption probability that is updated with time (every meal) according to the agent's socio-demographic characteristics, individual concerns, influence from other agents, and meat price. Agents in the same family can influence each other, and co-workers can be influenced through the workplace network [**?**]. This ABM approach to customer behaviour in meat consumption provides an idea to implement our ABM model. We could use a similar method to define our agents and market, extending from the demand for meat to the demand for other daily necessities to capture customers' spending habits. Aurelija et al. also conduct an ABM framework to investigate the customer behaviour in the insurance field that we can learn from the model implementation procedure [**?**].

Moreover, David Andersson developed an application to calculate the individual carbon footprint from the financial transaction data [**?**]. It used a novel hybrid approach to estimate the carbon footprint of users with three data sources shown in figure **??**. All transaction data are classified based on the Classification of Individual Consumption According to Purpose (COICOP), so the application could estimate GHG emissions according to different consumption categories of transactions [**?**].

In terms of agent-based modelling with influence strategies, Haer et al. present an example with an agent-based model to analyse flood risk communication strategies [**?**]. The paper analyses the effect on individuals with different flood risk communication strategies and the role of social networks in this process. The ABM model uses four different flood risk communication strategies compared to a baseline scenario with different probability and stander deviation to change agents' attitudes toward the flood risk [**?**]. Hence, we could use a similar approach to evaluate the most effective strategy that affects individual spending habits on greener companies.

By combining the related research above, we could have a clearer understanding of approaching the project step by step.

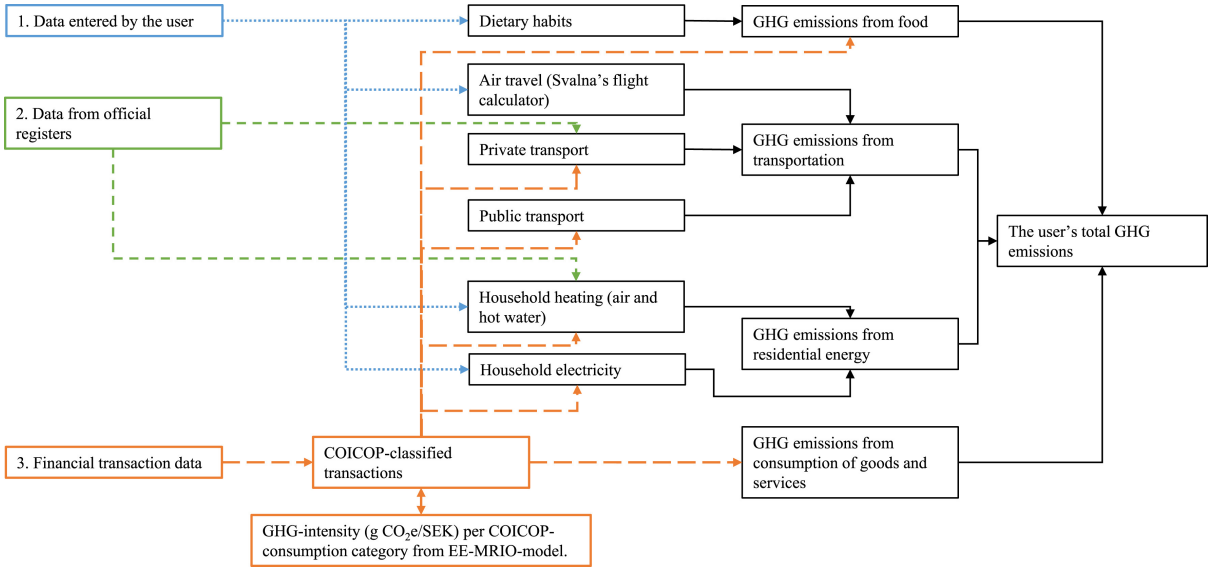

Figure 1: Schematic overview of various sources in order to estimate GHG emissions at the individual level

3 Programme and Methodology

3.1 Transaction Data

Bank transaction data are provided by ekko and are expected to contain basic payment information shown in table **??**. attributes of bank transaction data We also expected to have basic user information, including gender, year of birth, home location, average salary and other information that could help with analysis.

Attribute	Description
tran_ID	transaction ID
user_ID	user ID
acc_num	account number
tran_tp	transaction type (eg. credit, debit)
tran_date	transaction date
tran_cat	category of the transaction provided by bank (eg. supermarket, restaurant)
tran_amt	amount of money of the transaction

Table 1: Attributes of Bank Transaction Data

3.2 Data analysis

After obtaining and organising the project data, we could start data analysis using network science. To adjust and refine the ABM model, we need to find the pattern of the user groups and transaction data as the actual value. Moreover, we would like to identify the factors that affect spending habits with the user information. For example, place of residence has a substantial impact on people's spending habits based on exploring the individual economic behaviour [**?**]. We could classify the users into different groups according to the factors we

discovered, and groups can overlap. In addition, spending behaviour can also be characterised from the bank transaction data such as overall spending behaviour, temporal spending behaviour [**?**]. According to the classified user groups and characterised spending behaviour, we could draw a user sample model representing our unbiased ABM human agents. In the user sample model, each user has one or two characterised spending behaviour and several classified factor groups.

3.3 Model creation and evaluation

To build the ABM, we could reproduce the user sample model we analysed as human agents and form appropriate rules according to their characteristics. The update of the agent status is expected to be daily, and we could observe the result of agents' behaviour at daily intervals. The environment of the ABM should simulate the real-world market, which reflects the amount of carbon emission and commodity price. Different data sources can be used to modify the market and carbon emission, and multiple aspects can influence the real-world market. Thus, we could find a suitable market data combination according to the data pattern we obtained from the analysis. Validation of the ABM model is also based on the information we analysed from the actual data to confirm that we successfully simulated the user spending habits. Then, we could apply sensitivity analysis to analyse the model behaviour and output to ensure the model is valid.

3.4 Influence Strategy and Carbon Footprint Reduction

An influence strategy must be a reasonable strategy that can be applied to the real world and is worth comparing to the disburse. Therefore, we have to collaborate with ekko to discuss a couple of practical influence strategies. If available, we could also get some feedback or data on these influence strategies to help us apply the influence strategies to the ABM. Otherwise, we could make assumptions with ekko combining similar influence strategies results from other experiments to apply the strategies. There are also some network influence models that can be considered to model how information spreads through agents, such as the independent cascade model [**?**]. Marten's paper presents a method from a new perspective to target customers using the pseudo-social network that links the customers that transfer money to the same entities [**?**]. According to the results of different influence strategies we applied to the ABM, we could uncover the influence strategy or influence strategy combination with the best performance that encourages people to reduce their daily carbon footprint.

3.5 Risk Assessment

This project is more practical than theoretical, so there will be many unforeseen problems. As we can not obtain the data until the start of the program, there is the probability that the data is not what we expected or not available. Moreover, we do not know the selection procedure of the data yet. Although we assume that the users and their transaction data are selected randomly, bias might be involved. We do not know whether ekko will provide their carbon footprint calculation method, so we assume we would use transaction data to estimate the carbon footprint. Besides the main risks mentioned above, other risks are listed in table **??**.

Risk	Probability	Severity	Actions to Minimise Risk
Cannot get transaction data	Very unlikely	Severe	Find other available data sources as soon as possible
Cannot get transaction data on time	Unlikely	Moderate	Keep communication with ekko, ensure the latest time for the data and start the reading part of the project first
Cannot get user information	Likely	Moderate	Estimate the user group data with the open data source (e.g. Worldbank Database)
Data selection is biased	Likely	Minor	Although it would affect the analysis result, the ABM is also built with the biased data. Hence, the influence strategy is still suitable for the original users provided, and we would identify bias in the final report
No significant pattern while analysing transaction data	Likely	Minor	Try to identity hubs, community and any outlier that affect the result, combine the information we have to describe the data structure
Too much data to process	Unlikely	Not Significant	Separate the data according to month, weeks or day to analysis in time series, or separate the data according to user group
No significant spending behaviour difference	Unlikely	Minor	Simplify the characterised spending behaviour section
Cannot build a valid ABM simulate the given data	Unlikely	Severe	Revise the procedure of building the ABM, find out the possible problem, keep communication with supervisor to ensure the project is in the right direction
Cannot have influence strategies data	Likely	Minor	We could make assumptions and use similar experiment data. However, the accuracy of the result would be affected if applied to the actual situation
Leak user data	Very Unlikely	Severe	Find data leakage channels and stop losses in time. Try to keep the data locally without uploading it to other platforms, use a self-owned computer to complete projects and keep the computer safe

Table 2: Risk Assignment Table

3.6 Ethics

The project will abide by the ethical standards of the University of Edinburgh throughout. Although we would deal with the real user transaction data, we do not collect the data personally and have contact with any user. Thus, we do not need to apply for ethical approval for the project. Indeed, we still need to ensure user data security and prevent user information leaks mentioned in the risk assessment section.

4 Evaluation

The project data is provided by ekko, so we do not have problems related to data collection. To ensure the ABM built is valid, we would conduct the model validation using analysed results and sensitivity analysis mentioned in the model evaluation section. Influence strategy selection is based on the result of influence on the valid ABM model. Although we might need to make some assumptions about the impact of the strategies, the assumptions would be made based on experiments and data analysis. We would run validation tests or prove with data on the assumptions created. Thus, the influence strategies selected are valid and unbiased if the original data provided is unbiased.

5 Expected Outcomes

To conclude, the project would present an agent-based model that simulates the user spending behaviour analysed from the transaction data provided by ekko. Moreover, we would find out the influence strategies that maximise the influence on encouraging users to purchase at sustainable companies. This project will deliver a different point of view in the analysis of daily spending behaviour as it uses the ABM approach, which is not common. Moreover, our ecosystem is threatened by the climate change problem, and each individual needs to take responsibility. The project could be the bond between reducing carbon emissions and daily life. Although this project might be less helpful in filling theoretical gaps, it is practical and can be put on the market after certain considerations.

6 Research Plan, Milestones and Deliverables

The project can be divided into three stages: preparation, implementation and documentation, and the implementation of the project is split into four parts. The Gantt chart demonstrates the project plan shown in figure **??**. The black lines in the Gantt chart indicate the expected time for the stage, and the blue bar indicates the time predicted for each part of the stage. The black lines are slightly longer than the blue bars in each stage as we set aside time to deal with unexpected situations. The first stage is the preparation which we wait for the industry data and complete the related reading. The data is expected to be delivered in early June, so we expect to finish the reading while waiting for the data and organise the data by the middle of June. We could start the data analysis process once the data is organised, and we expected to spend eight to ten days for each part in the implementation stage. Data analysis and model evaluation might take less time than others, and influence strategy might need to communicate with ekko, which increases the estimated time. The planned time is relatively generous to ensure each part can be completed on time. There are two documents that need to be delivered

during the project procedure. Project progress is required in week 6 with a short length so that we could plan less time on it. The dissertation is the most crucial part of the project with a heavy workload. Therefore, we could start writing along with the project progress and refine the dissertation after project implementation is completed. Milestones are listed in table **??** with a two-week interval to ensure the stability of the project schedule. Deliverables are listed in table **??** according to the milestones.

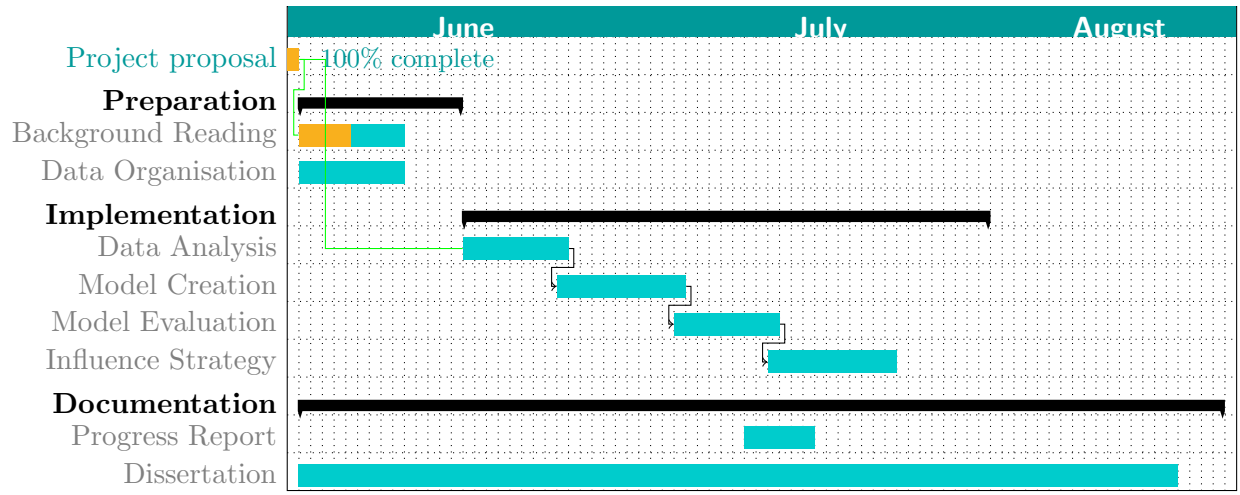

Figure 2: Gantt Chart of the activities defined for this project.

Milestone	Week	Description
M_1	2	Preparation completed
M_2	4	Data analysis completed
M_3	6	Model creation completed
M_4	7	Model evaluation completed
M_5	8	Project implementation completed
M_6	10	Submission of dissertation

Table 3: Milestones defined in this project.

Deliverable	Week	Description
D_1	4	Data analysis result
D_2	6	ABM model
D_3	6	Progress report
D_4	8	Influence Strategies
D_5	10	Dissertation

Table 4: List of deliverables defined in this project.

References

[1] Footprintnetwork.org. 2022. Climate Change & the Carbon Footprint. [online] Available at: ¡https://www.footprintnetwork.org/our-work/climate-change/¿ [Accessed 27 April 2022].

[2] Selin, N., 2020. carbon footprint — Definition, Examples, Calculation, Effects, & Facts. [online] Encyclopedia Britannica. Available at: ¡https://www.britannica.com/science/carbon-footprint¿ [Accessed 27 April 2022].

[3] Bonabeau, E., 2002. Agent-based modeling: Methods and techniques for simulating human systems. [online] Available at: ¡https://www.pnas.org/doi/10.1073/pnas.082080899¿ [Accessed 27 April 2022].

[4] North, M. and Macal, C., 2007. Managing business complexity : b discovering strategic solutions with agent-based modeling and simulation. Oxford: Oxford University Press, p.328.

[5] Barabási, A., 2013. Network science. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 371(1987), p.20120375.

[6] Menczer, F., Fortunato, S. and Davis, C., 2020. A First course in network science. Cambridge [etc.]: Cambridge University Press.

[7] Rand, W. and Rust, R., 2011. Agent-based modeling in marketing: Guidelines for rigor. International Journal of Research in Marketing, 28(3), pp.181-193.

[8] Scalco, A., Macdiarmid, J., Craig, T., Whybrow, S. and Horgan, G., 2019. An Agent-Based Model to Simulate Meat Consumption Behaviour of Consumers in Britain. Journal of Artificial Societies and Social Simulation, 22(4).

[9] Ulbinaitė, A. and Moullec, Y., 2010. TOWARDS AN ABM-BASED FRAMEWORK FOR INVESTIGATING CONSUMER BEHAVIOUR IN THE INSURANCE INDUSTRY. Ekonomika, 89(2), pp.95-110.

[10] Andersson, D., 2020. A novel approach to calculate individuals' carbon footprints using financial transaction data – App development and design. Journal of Cleaner Production, 256, p.120396.

[11] Haer, T., Botzen, W. and Aerts, J., 2016. The effectiveness of flood risk communication strategies and the influence of social networks—Insights from an agent-based model. Environmental Science & Policy, 60, pp.44-52.

[12] Sobolevsky, S., Sitko, I., Tachet des Combes, R., Hawelka, B., Murillo Arias, J. and Ratti, C., 2016. Cities through the Prism of People's Spending Behavior. PLOS ONE, 11(2), p.e0146291.

[13] Tovanich, N., Centellegher, S., Bennacer Seghouani, N., Gladstone, J., Matz, S. and Lepri, B., 2021. Inferring psychological traits from spending categories and dynamic consumption patterns. EPJ Data Science, 10(1).

[14] Martens, David and Provost, Foster, Pseudo-Social Network Targeting from Consumer Transaction Data (September 2011). NYU Working Paper No. CEDER-11-05, Available at SSRN: https://ssrn.com/abstract=1934670