Identifying Top Influencers in a Social network

Opinion Dynamics within Online Communities

Presented to: Scott Kyle

Serdar Yüksel, and Chris Bentz

APSC 200

Applied Mathematics

Faculty of Applied Science

Queen's University

Prepared By: Team 5, Section 208

Maeve Buchan: 20020510

Cooper Midroni: 20002125

Sean Kato: 20012208

Noah Gutmann: 20028250

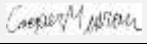
December 3rd, 2017

Statement of Originality

Our signatures below attest that this submission is our original work.

Following professional engineering practice, we bear the burden of proof for original work. We have read the Policy on Academic Integrity posted on the Faculty of Engineering and Applied Science web site (http://engineering.queensu.ca/policy/Honesty.html) and confirm that this work is in accordance with the Policy. http://engineering.queensu.ca/policy/Honesty.html) and confirm that this work is in accordance with the Policy.

Table 1: Statement of Originality signatures.

Name: Maeve Buchan Signature: 	ID: 20020510 December 3rd, 2017
Name: Sean Kato Signature: 	ID: 20012208 December 3rd, 2017
Name: Noah Gutmann Signature: 	ID: 20028250 December 3rd, 2017
Name: Cooper Midroni Signature: 	ID: 20002125 December 3rd, 2017

Space below is for a table showing which portions of this report each individual made a significant contribution toward.

Table 2: Work Distribution.

	Content	Write-up	Editing
Cooper Midroni	Problem Statement, Problem Scope, Stakeholders, Design Criteria, Final Design Solution, Concept Selection, Assessment of Engineering Specifications.	Problem Statement, Problem Scope, Stakeholders, Design Criteria, Final Design Solution, Concept Selection, Assessment of Engineering Specifications, Recommendations, Assessment of Design Criteria.	Whole document
Sean Kato	Background Research, Stakeholders, Project Economics, Assessment of Design Criteria, Recommendations, Conclusion.	Background Research, Stakeholders, Project Economics, Recommendations, Assessment of Design Criteria.	Whole document
Maeve Buchan	Executive Summary, Project Management, Background research, Social/Environmental Considerations, Concept Selection, Assessment, Conclusion	Executive Summary, Social and Environmental Considerations, Ethical Concerns, Assessment of Design Criteria.	Whole document Document Formatting
Noah Gutmann	Background Research, Project Economics, Safety Considerations, Conclusion	Introduction, Project Economics, Safety/Regulatory Considerations, Conclusions	Whole document References

Executive Summary

As the world becomes increasingly digitalized, companies have turned to social media as a platform for marketing. Social media websites now account for 30% of today's internet use [2], making the more popular sites an effective way to spread a message, idea, or product. The objective of this model is to determine the ideal agents for advertisement targeting within a system, using both opinion dynamics and optimal deployment. The model will be constructed for both general and specific use. It will have the ability to portray any scenario given the necessary input data, as well as more general settings for observational analysis.

The stakeholders considered in this project include marketing agencies, high and low salience companies, consumers, and owners of data. Using stakeholder needs as a guideline, a list of customer requirements was developed. These requirements are: scalability, accuracy, optimization, increase in saliency, data security, and cost-effectiveness. A weight was assigned to each requirement, which was applied in the Quality Function Deployment (QFD) to evaluate the sensitivity of engineering specifications.

Following the QFD, an evaluation matrix was used to select the optimal design solution for each design characteristic. Of the considered characteristics, the final design solution concluded to use file extraction as the method of data input, and a manual input to specify ad location in the model. To enhance model accuracy, trust values will exist between the ad and the target agent, as well as between all agents within a given sphere of influence. Finally, a complete iteration will score each agent based on their proximity to the advertisement. The highest score defines the top influencer of the system.

In order to recognize these 'ideal agents', several areas must be considered. An agent's sphere of influence considers the power, legitimacy, and urgency of their social media presence. This, along with data received from internet cookies and data providers can inform stakeholders about individual consumer habits. This data will be used to generate a range of trust values, as well as to place agents on the opinion plane.

As the model utilizes personal information, there are social and ethical considerations that must be acknowledged. As the model optimizes on opinion-based campaigns, it has the ability to strongly impact the spreading of an idea. Thus, it is important to ensure the intentions of an ad campaign before the idea is set to spread. This level of consideration also applies to the security of the data being used. It is vital that information be encrypted and stored through proven shard-ing methods.

With regards to project economics, research concludes that pricing for online marketing is typically based on the number of clicks an advertisement receives. To acquire the proper data, a typical sum for the appropriate sets of data is $0.05 per. To run the model analyzing a group of 20 agents, the associated cost would be $2.72 assuming the ad is only clicked once.

The model in its current capacity has many simplifications within the project scope. As the model becomes more advanced it will grow to represent more opinion dimensions to better predict the complexity of human behaviour. The model should expand to account for time-variable trust values and ad positions, enabling the model to project ad campaigns in the long term. With these developments comes the opportunity to account for a competitive market space, all contributing to a drastic increase in accuracy.

Table of Contents

List of Figures

List of Tables

1.0 Introduction

1.1 Background Information

A high-salience company is defined as a company with notable prominence in the market and in the minds of consumers. Three aspects can measure this concept of salience: power, urgency, and legitimacy [1]. Each aspect was given a score of either low, medium or high. A weighted average of these aspects can determine the total salience of an agent. Power is defined as the trust value between agents, and legitimacy the relevance to each agent. For simplicity, legitimacy was fixed at a high value for all agents. Urgency is the location of an agent with respect to the target location.

The concept of defining and determining the saliency of individual agents can be extended towards the determination of an agent's range and impact of influence in this model. Similar tests could be performed, specifically on social media users, to determine values such as online engagement, and interaction with others. Influence on social media may be drawn from any number of factors, such as: quantity of friends/followers, mutual connections/interests, hours logged (how many and at what time), and responsiveness from surrounding agents. An agent with far reach but low responsiveness may be an example of a lower 'power rating'. This is an important measurement, as an agent that reaches a smaller core group with high idea reception can ultimately have more influence than the converse. The abundance and type of data available to this project will greatly affect what aspects of saliency will be defined in further stages of the report.

The notion of 'trust values' between agents is similarly important. Trust values indicate how permeable an agent is to the ideas of its surrounding agents, and they are assigned per individual connection. One way to approximate trust values between agents is to use a trust metric. The most accepted way of determining trust values is using an empirical metric [2]. Empirical metrics combine observed human interactions/behavior, and theoretical background research to hypothesize expected behavior [2]. The value of the levels of trust is determined to be the difference between these observed and expected behaviors [2].

Observed levels of trust are often gathered by surveys. One of the most well-accepted means of gathering unbiased social opinions from the masses is the General Social Survey (GSS) [3]. The GSS covers a wide range of topics, the most notable to this project are behavioral and attitudinal questions [3]. These questions allow people to anonymously discuss their thoughts and feelings in the context of trust and cooperativeness, allowing scholars to analyze and place trust values on these answers and percentages [4]. Another advantage of the GSS is that it gathers and sorts data based on the

demographic of the users. In other words, it allows for easy analysis of trust values associated with different age brackets and ethnicities [4]. As this was just one example of a survey, many others have been created to gather more specific results. For this project, trust values will be gathered from these existing surveys and through the generation of an entirely new survey/poll that will be accessible to agents through Facebook.

To place agents accurately on a desired opinion plane, it is necessary to acquire the proper data. The most common use of data with relation to this project is known as targeted advertisement. Companies and websites which are able to collect a wide range of information (search history, purchase habits, etc.) are able to make informed decisions regarding who sees their advertisements; they do this in an attempt to increase ad engagement. Engagement is a term used to indicate the likelihood of consumer interaction with the ad (clicking through, watching a video for longer, sharing, etc.).

The method by which companies and websites receive personal data is through internet cookies, a file that contains individual records of all previously browsed web pages. There are two types of internet cookies: First-party, which are sent by the site domain, and third-party, which are stored and come from other site domains [5]. First-party data includes the frequency of website visits, time spent on a webpage, which ads have been previously clicked on, and content that the user has downloaded [6]. Third-party data can be bought from other sources, and can include information such as user demographics and a consumer's online purchase history. From the information collected through first-party cookies, ad agencies can learn about consumer's interest and behaviors [6].

Other data platforms are additionally used to better understand consumer interest; these include: search data, purchase data, and profile data. Search data analyzes user habits and searches to determine which targeted ads will be shown beside search results [5]. Purchase data recommends products based on past purchases made on other online stores. Profile data chooses custom ads based on personal information from online profiles [5]. The information collection can include: age, religion, education, political views, and favourite entertainment genres.

1.2 Problem Statement

It is estimated that presently, 30% of all internet use is spent on social media interaction, with certain demographics spending up to nine hours a day on social platforms such as Facebook, Twitter, and Snapchat [7]. This fact represents a drastic reallocation of the average population's time and attention. To adjust to this shift, the practice of commercial marketing has applied itself to online advertisements. To better focus marketing campaigns, it is critical to utilize available data to optimize the impact of

advertisements on online communities. The model will utilize gathered information to determine an optimal target agent which in turn influences others towards a preselected opinion point. This will maximize the net gain of a company's advertising efforts.

1.3 Project Scope

This project aims to create a small-scale model of an online social network, consisting of between 10-20 agents. It will consider the sphere of influence of individual agents, as well as the associated trust values between agents in communication with one another. Its primary purpose will be to identify optimal agents within the system, which when targeted, will provide the greatest desired impact on the entire system. The project will remain generic in its creation of an opinion plane, and will not seek to provide insight regarding specific social topics. For simplicity the model will consider a two-dimensional opinion plane depicting four different stances on any given product or ideology. The four stances will remain generic for this model, allowing it to be applied for any given situation.

This generic model is justified given the fact that the purpose of the model is to determine the optimal agent to increase the return on investment of a company. Thus, it will not be hand-tailored to a specific product.

2.0 Discussion

2.1 Stakeholders

The creation of this model has the potential to impact a diverse range of stakeholders. Notably, key stakeholders include advertising agencies, high and low salience companies, owners of data, and consumers.

The most immediate stakeholders of this project are advertising agencies, specifically those utilizing the model itself. The model provides advertising groups the ability to improve current performance on their ad campaigns. It will achieve this feat by maximizing the distribution and influence of their ads using available data. Advertising companies will require the model to be easily scalable, to adhere to a market of varying sample size and dimension. This implies that generality will be an important feature to the model.

It is considered stakeholders that are considered 'clients' to the project will benefit directly from the success of the model. This includes corporations attempting to use online advertising to increase the success of their business. Companies that use online marketing can be separated into two categories: high-salience and low-salience organizations. The reason for this distinction is that high-salience

companies are substantially less reliant on social media for advertising. They are able to partake in mass advertising campaigns, and can utilize large marketing budgets to reach their consumers. Low-salience companies are much more reliant on online marketing, as it is one of the cheapest forms available. Increasing online presence in the eyes of several thousand consumers is significantly more impactful to a start-up than for an organization such as Coca-Cola. For this reason, high-salience companies have a considerably smaller stake in this project compared to low-salience companies. Any organization hiring these ad agencies would require that the model be cost-effective, as well as reasonably accurate for their specific market of consumers. These stakeholders would require that the model on ad efficiency, and ultimately be able to recommend a method of marketing which achieves more salience per dollar.

Owners of data represent a significant, and incredibly vital stakeholder to this project. These are organizations such as Facebook, Google, or Amazon, which can track individual user statistics over extensive online use. These organizations will provide the data which will inform the model, and ultimately place agents in their appropriate space on the opinion plane. It is important to note that this stakeholder group is not limited to massive social networks, but also smaller companies that are able to provide useful, perhaps even more insightful data, through niche use of an application or product. Thus, these organizations will likely pursue some form of partnership, and will require that the model be able to process their data effectively and safely. This stakeholder will also require well defined dimensions of the opinion plane, to understand what data will be useful for the model.

The remaining prevalent stakeholders of this project are the consumers. Consumers can be further divided into primary and secondary consumers. Primary consumers are the determined ‘optimal agents’ and thus, will be directly exposed to advertising, as dictated by the model. These consumers will have the greatest net effect on their surrounding network, and will be the ones which influence the secondary consumers. As primary consumers interact with the advertising, they will require that any advert be ethical, tasteful, and promote a positive impact. This provokes a strong conversation about the ethics of ‘targeted advertising’, and the use of personal data. It will be necessary to preserve the confidentiality of consumer information, and that it be collected ethically. Secondary consumers are merely a by-product of the radius of confidence of primary consumers, resulting in having a relatively low stake in the project itself. Secondary consumers must be considered, as when influencing any opinion, it should be noted that the marketed ‘idea’ is ethical and just.

2.2 Design Criteria and Specifications

A crucial aspect of this model will be its generality, and as a result it's flexibility. The social network model must have the ability to apply itself to any opinion plane. To satisfy this, the axes will simply be referred to as 'X' and 'Y' until applied to a specific opinion scenario. A feature of the 'general model' must be its ability to take user input, and create the consequent space within the model. Areas of user input will consider: number of agents, type of distribution (random, even, polarized, etc.), and area of focus (the desired area for agents to converge). This will allow for demonstrations to aid in user understanding of the model's use and function.

To ensure the model is functional, it is necessary that it be able to portray specific network scenarios with no limitations. Information regarding specific scenarios may be uploaded through a Microsoft excel file. Therefore, it is a requirement that this information can be processed through MATLAB to build the opinion-scape. An uploaded file will contain information such as: total quantity of agents, plane location, data to calculate the area of influence, and associated trust values.

It is also required that this project have the ability to scale. A base space of 10-20 agents will be considered for a general case, though the model should have the ability to iterate through up to 100 agents. Considering this, the model's iteration must be efficient when identifying the top influencer of the system. As such, it is a requirement to ensure that the model meets standards in iteration, and that the method of processing is as non-exhaustive as possible. It will be necessary to research the ideal method of sorting for the model's particular problem. A strong measure of the complexity of the iteration will be calculating the 'Big O' execution time [8]. It is similarly necessary to determine a method of 'scoring' each iteration against the desired input. This will allow for the model to conclude which agent(s) influence the system to the greatest degree.

A major constraint on this model will be its inability to model a realistic number of opinion dimensions. There are an unquantifiable set of factors which determine an agent's stance on an opinion. As this model only sets out to consider a two or three-dimensional plane, a large portion of influencing factors will be ignored. This model will therefore operate under the assumption that the dimensions of the model are the only determining factors for agent opinion.

2.3 Environmental Factors

As the world conforms to a global shift in marketing mediums and strategies, an accompanying change in environmental impact arises. Through the increased the use of digital marketing and online advertisements, there is a natural reduction in the practice of implementing physical advertisements.

Among these include: newspapers, magazines, flyers, and even billboards, whose lack of use imply a decrease in total waste production. It is evident that digital marketing is a greener solution to paper and plastic alternatives, though it is important to still consider the wasteful habits that come hand in hand with this new marketing practice. This surge in digital marketing is, in fact, a by-product of an even greater societal trend; the mass adoption of technologies. In 2015, 68% of U.S adults reported owning a smartphone, compared to 33% in 2011, more than doubling in the span of 4 years [9]. As digital marketing is successful based on the notion that electronic products are the new 'medium of advertisement', the waste of these devices must be considered. It has been calculated that approximately 50 million tons of electronic waste are produced each year, with 3 million tons directly a result of the United States [10]. This number is only expected to rise as the digital world grows, and developing countries are brought into the prosperity of technology that the developed world inhabits.

This project allows for the opportunity to address a variety of environmentally oriented factors. Namely, this opportunity is attached to the notion that data collection could uncover significant social patterns relating to the environmental awareness of social groups. Along this spectrum there exist groups which are environmentally destructive or helpful, allowing for specific targeting for the purpose of social change. As the model can optimize the use of its algorithms to enhance environmental campaigns and advertisements, this provides a great boon to campaigns of activism and awareness. Companies such as PETA and Patagonia, which emphasize sustainability and environmental care, would be able to reach a more compatible audience for their business success. The contrary view is that this project could similarly be used in an attempt to alter the opinions of those who are considered un-environmental. This implication is present for any campaigns centering on pertinent social issues (i.e. politics, religion, etc.).

2.4 Social Considerations

Due to the nature of opinion altering, there are incredible social considerations that are crucial to acknowledge. It is first noted that this project is reliant upon the fact that an individual's opinions and behaviours are tracked and quantified by social media and data gathering organizations. This itself is an incredibly contentious point, as the internet grows at such a pace that legal policy surrounding the privacy of data often lags behind.

Using consumer data is, however, a logical choice for advertisement agencies due to its proven efficiency and abundance. Though data use carries with it the controversy of breaching privacy and ethics. The concept of being monitored and the fear of what companies may do with personal information is a legitimate concern for internet users [11]. This model creates the potential to further

manipulate agents into purchasing products, potentially compromising their economic or even emotional positions by influencing them to believe that it is in their interests to purchase products or support an idea.

It should be noted that model's algorithm can also be used to promote humanitarian projects and socially responsible companies more effectively than ever before. For example, at the occurrence of natural disasters there arises an opportunity to promote monetary support or even volunteer relief efforts that are pivotal in aiding recovery. By reaching out to the correct demographic, there is a greater likelihood of response and positive feedback.

Economically, this project could support new companies which are attempting to take off, to help them find their ideal audiences at a critical time in their development. Marketing to the wrong demographic can be a detrimental action for a company's legitimacy, thus supporting the importance of the model.

Another important social aspect to consider are the jobs that this model will create or perhaps destroy. In the past 3 years there has been a 73% increase of in the use of digital marketing [12]. Along with this new category of marketing comes new jobs in fields related to digital marketing. On the other end of the spectrum, owners of billboards and physical marketing products will see a decrease in interest for their service. This poses the risk to harm revenue and career opportunities for those working in these areas. Newspapers and magazines are sustained mainly through the generated income of advertising [13]. With a decrease in need for physical marketing, the probability of these services becoming obsolete increases.

2.5 Ethical Considerations

Similar to social considerations there are many ethical dilemmas that accompany this model. In order to effectively apply the algorithm, a diversity of data must be gathered which may include personal information on agents. Categories such as age, lists of friend and associates, spheres of influence, physical health and many more can be considered by the model. Collecting and monetizing this information is often seen as an unethical practice. Countless consumers are displeased that companies such as Facebook are documenting user activity. Thus, by capitalizing on the collected data, the project encounters a moral grey area.

It is widely acknowledged that in the foreseeable future the advent of high-accuracy probability algorithms will guide predictions on consumer behaviour, perhaps even ahead of the consumer's own decisions. Algorithms are able to make these predictions based on their mass accessibility to data, based on the aptitude of computers for pattern recognition. These behaviour patterns may be found in word

use, user location habits, and even a consumer's engagement with online content. To prevent a potential breach of privacy, a code of ethics to govern these algorithms should be implemented, with a focus on the impacts of online advertising. Considerations of what is acceptable must be directly outlined and approved in order by legislative bodies to reduce the likelihood of overstepping boundaries.

2.6 Safety Considerations, Regulatory Compliance, and Professional Practice

As the model uses personal agent data, it is important to guarantee data security, to prevent criminal or malicious acts from occurring. For this reason, the data must be stored and encrypted properly to prevent external parties from obtaining the information [14].

Generally, companies use online cloud services to store user information as they can securely encrypt the data to keep it safe [14]. This does, however, comes with the minor risk that an employee can access data from within the company. To avoid such a scenario, websites can offer an option for customers to be responsible for their own encryption, negating the risk of employee access [15]. Amazon uses a cloud storage process called 'shard-ing' by which they break each data file into multiple encrypted files, which are then stored in different places on the cloud. By this process, only fragments of user data can be obtained, lessening the impact of compromised information [15].

Most websites have their own authority to use data collected from their site. This occurs mainly due to clauses in the Terms of Service and Privacy Policy which are agreed to without being read [16]. In doing so, users often consent for their data be sold for marketing purposes. In 2017, Europe introduced a new law to govern online advertising, requiring that a website now ask a user whether they want to be tracked or not. In the past, users were automatically tracked [17]. The expectation is that most users will choose to not be tracked, making it harder for companies to use target advertising, as there will be less data available to collect. However, there are currently no laws that prevent companies from collecting information to use for online advertising and marketing [16].

3.0 Concept Selection

3.1 Assessment of Engineering Specification Sensitivity

To guide the preliminary concept selection, a Quality Function Deployment (QFD) was performed. The QFD was used to identify the engineering specifications of the proposed opinion model which are most sensitive to client needs. This information was used to create a wider variety of design solutions on topics that were most sensitive, so as to better allocate team resources.

The process began with a period of time in which the potential needs of each unique stakeholder was contemplated. This resulted in a complete list of all customer requirements. Once compiled, similar needs were combined into more inclusive criteria, allowing for the needs of each stakeholder to be well represented. The results of the customer requirement analysis can be seen below in Table 3.

Table 3: A table of the final customer requirements to be analyzed in the QFD.

Customer Requirements	
Scalable	Optimized Output
Adaptable	Increase to Saliency
Accurate	Security of Information
Cost-Effective	Ethical Data Use
Non-Intrusive	

Following the completion of the customer requirements, it was then necessary to assign a weighting to each criterion. To maintain stakeholder representation, the importance of each criteria was evaluated from the perspective of each stakeholder. These stakeholders are seen below in Table 4.

Table 4: A table of the major stakeholders of this project.

Major Stakeholders	
Ad Agencies	Owners of Data
High Salience Companies	Consumers
Low Salience Companies	

To view the assigned weights, refer to

Appendix A: Quality Function Deployment All stakeholders in this project were held equal in the conclusion of weight, in which the value for each criterion was averaged. Each stakeholder was held equal on the reasoning that they all share parallel interests in the project. It can reasonably be seen that advertising agencies are as reliant on the success and accuracy of the project as are the companies that the model will be used for. Ad agencies would desire a scalable model, such that it can conform to varying target market size. This scalability would benefit both high and low salience companies on the premise that they may choose any audience to preview their product to. With regards to the ethical concerns behind data use, it is similarly impossible to separate the concerns that large data owners have with those of the consumers they represent. For both parties, it is pivotal to maintain the security and safety of data as it is passed between data owners to the model. For the agencies and companies that profit from the project, an unethical use of data may be grounds for a lawsuit. Thus, the criteria are seen to hold a shared importance for each stakeholder. As such, the combined weightings were simply averaged, and all stakeholders held to the same rank of importance. The final weightings are determined (to one decimal) for each customer requirement as seen in Table 5.

Table 5: Holds the final weights of each customer requirement to be used in the QFD.

Customer Requirement	**Final Weighting**
Scalable	17.7
Adaptable	17
Accurate	17.8
Cost-Effective	15.7
Optimized Output	17.7
Increase to Saliency	14.9
Security of Information	16.4
Ethical Data Use	13.4
Non-Intrusive Advertising	11.4

With final weights decided, the following engineering specifications were analyzed in the QFD.

Table 6: A list of the engineering specifications to be analyzed under the QFD.

Engineering Specifications	
Method of Data Input	Agent Interaction
Ad Interaction	Agent Iteration
Scoring Algorithm	Ergonomic Features

The specifications above were selected to, through the QFD, better understand the importance of each specification to the stakeholders of this project. Look to Table 7 below for a justification as to why each specification was chosen.

Table 7: Summarizes the justification for selecting engineering specifications.

Engineering Specification	Justification (To evaluate the importance of…)
Method of Data Input	How data is brought within the model. Whether it is ideal to use automated processing, or to include a degree of human intervention.
Ad Interaction	How the advertisement interacts with agents. What degree of trust does an agent have with an ad? Is there a limited range that ad can operate within?
Scoring Algorithm	The method used to determine the success of ad targeting. The way success is scored will affect which agents are selected as top influencers.
Agent Interaction	How agents interact with one another, whether through trust values or assumed blind trust.
Agent Iteration	Identifying an ideal method to iterate through each agent, in determination of the most influential agent in the system.
Ergonomic Features	Implementing methods which allow the model to become more interactive, and perhaps yield more insight.

The completed quality function deployment can be seen in

Appendix A: Quality Function Deployment. The results can be seen below in Table 8, where the sensitivity of each engineering specification is indicated by its associated QFD value. A higher value indicates that the specification is more sensitive to customer needs.

Table 8: Summarizes the results of the QFD.

Quality Function Deployment Results			
Engineering Specification	**QFD Value**	**Engineering Specification**	**QFD Value**
Method of Data Input	744	Agent Interaction	538
Scoring Algorithm	734	Ad Interaction	782
Ergonomic Features	251	Agent Iteration	413

3.2 Selection of Design Solutions

Guided by the sensitivity values of the QFD, the quantity of design solutions seen in Table 9 is representative of stakeholder needs. Those specifications with less solutions were determined to be less important based on the customer requirements of the project.

Table 9: Displays the chosen design solutions for each engineering specification.

Engineering Specification	Design Solution #1	Design Solution #2	Design Solution #3
Ad Interaction	Limited Range	Ad Trust Values	Ad Trajectory
Method of Data Input	File Extract	Manual Input	Interpolate Data
Scoring Algorithm	Ad Radius	Agent Convergence	Net Change
Agent Interaction	Agent Trust Values	Equal Consideration	
Agent Iteration	Complete Iteration	Sorted Iteration	
Ergonomic Features	Ad Location Input	Ad Trust Value Slider	

The list of customer requirements was compared with the above design solutions in an evaluation matrix to determine the ideal solution for each engineering specification. Possible solutions were rated on a scale from 1 to 4 with reference to their ability to satisfy the corresponding customer requirement, and then multiplied according to the requirement's weight. The sum of these weighted values represent the implied value of the design solution, with higher values being deemed superior solutions given stakeholder needs. The results of the evaluation matrix can be seen in

Table 10 below.

Table 10: Displays the evaluation matrix results for each design solution.

Engineering Specification	Design Solution	Evaluation Matrix Total
Method of Data Input	File Extract	**348**
	Manual Input	279
	Interpolate Data	314
Ad Interaction	Limited Range	302
	Ad Trust Values	**324**
	Ad Trajectory	278
Scoring Algorithm	Ad Radius	**432**
	Convergence of Agents	334
Agent Interaction	Net Change	290
	Agent Trust Values	**327**
Agent Iteration	Equal Consideration	302
	Complete Iteration	**350**
Ergonomic Features	Sorted Iteration	328
	Ad Location Input	**356**
	Ad Trust Slider	298

Bolded values in

Table 10 represent that the corresponding design solution was scored the highest in the evaluation matrix. For more details regarding the weighting and scoring of each design solution refer to Appendix B: Evaluation Matrix.

4.0 Final Design Solution

Table 11: Displays the most fitting design solutions, as selected by the evaluation matrix.

Engineering Specification	Design Solution
Method of Data Input	File Extract
Ad Interaction	Ad Trust Values
Scoring Algorithm	Ad Radius
Agent Interaction	Agent Trust Values
Agent Iteration	Complete Iteration
Ergonomic Features	Ad Location Input

Table 11 above displays the design solutions as selected by the evaluation matrix. Of the six engineering specifications this report will distinguish the categories of Method of Data Input and Ergonomic Features

as the ‘Interface Design’ component to the project. These specifications were chosen as they relate to how users will interact with the model, by way of preparing data for input and adding an element of human control. The other four specifications, Ad Interaction, Scoring Algorithm, Agent Interaction, and Agent Iteration, will be referred to as the ‘Model Design’, as they consider the mathematics that will be at work behind the user interface.

4.1 Interface Design

Method of Data Input

In the determination of the method of data input, all customer requirements were considered, with some being more relevant to the engineering specification than others. For instance, the method by which data is taken in by the model has little to do with the intrusiveness of an advertisement, but vastly affects the data security of the model. It was determined by the customer criteria, and by objective evaluations contained in the rubric, that extracting data from a given file was the optimal solution.

Through this method, files can be encrypted and sent over secure networks, which comes with a variety of benefits. The use of encryption encourages data to be moved to the cloud [18]. The cloud is a more flexible and low-cost form of storage, and the use of encryption keys guarantee the access of data to a select few. Furthermore, using encryption will satisfy regulations such as the Personal Information Protection and Electronic Documents Act [19], which serves to legitimize the use of the model as a whole, and can be used to gain access to more insightful data. Additional benefits of file extraction coupled with encryption includes the ability to safely back up files, and to ensure that geographically separate offices are secure, allowing the model to be accessible worldwide [18].

Ergonomic Features

It was determined that the ideal solution for an ergonomic feature in the model would be to allow the manual input of the advertisement location. This solution aims to provide key insight into effective locations for the advertisement. An article titled the Principles of User Interface Design promotes the thought that the purpose of interface is to enable interaction and to keep users in control [20]. By allowing the manual input of the advertisement location, customers are able to choose the proximity of their advertisements to their target audience. This insight can contribute to the decision of a company pursuing long term or short-term marketing strategies. Short term marketing is ideal for increasing online conversions (turning ad placements into website views, successful sales, etc.), while long term marketing focusses more on content driven ad placement [21].

As the model works on the basis of iteration, it is natural to extrapolate each iteration to represent a given time step. As such, a company can choose their campaign's purpose to increase website views over the span of a week, or to attempt to develop a more natural viewership over months. Using the manual ad location input, the effectiveness of individual advertising can be perceived at will. As such, with a properly modeled space, any stakeholder can use this insight to determine their best course of action. The user interface containing the ad location input can be seen below in Figure 1.

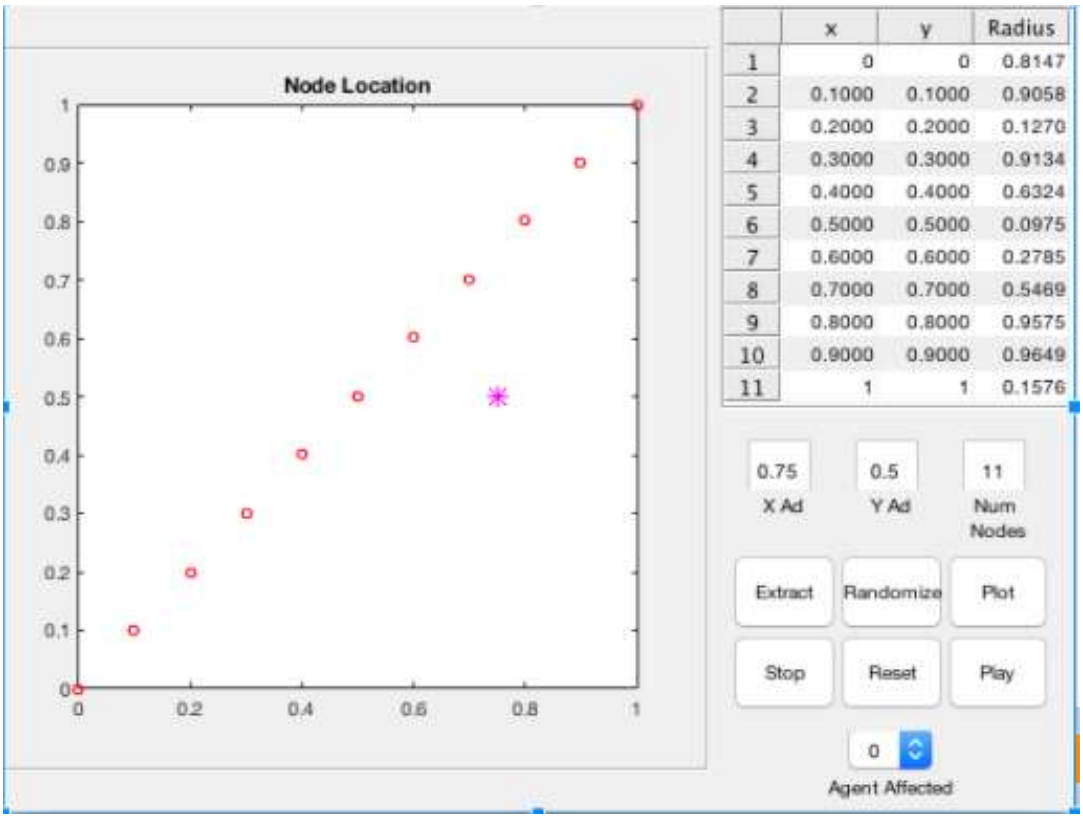


Figure 1: Displays the user interface of the model, including the opportunity to specify the ad location.

4.2 Model Design

Ad Interaction

It was selected that the ideal solution as to how advertisements interact with individual agents is to implement a trust value between the advertisement and the target agent. This enhances the accuracy of the model, as the assumption that agents are entirely responsive to ads is inherently false. A study was conducted on the degree of trust between mobile advertisements and their target audience. It was concluded that user trust as well as the target's degree of familiarity with the marketed topic can determine 54% of their change in attitude towards the advertisement [22]. This infers, among other insights, that it is possible to use agent data to determine the associated ad trust value. Similarly, in 2013 it was measured that trust in online banner ads jumped to 42% from 26% in 2007 [23]. Thus it becomes pertinent to utilize available data to optimize on agents which are more trusting of the advertisement topic.

Scoring Algorithm

The mathematical process for how 'success' is determined in this project is the most influential engineering specification for the model as a whole. The particular approach to scoring the opinion space is in essence the main deliverable of the project; a measure for how the model predicts ads to perform. Thus, it becomes pivotal to produce quantitative values to act as a measure of how favourably the opinion space is influenced. This algorithm must realistically simulate how purchasing habits are altered by an agent's location on the opinion plane, and the degree to which it aligns with a favourable outcome.

According to a variety of psychological studies, as one would expect, altering the opinions of individuals who have particularly strong beliefs is incredibly more challenging than those who hold loosely to their opinions [24]. It even describes that those individuals who hold the strongest beliefs can tend to react negatively when shown contrary opinions. In Krause mathematics, this can be thought of as an agent moving further away on the opinion from the location of the advert. A rigorous study in the UK produced the following chart in **Error! Reference source not found.**, depicting the converging of 100 unique opinions over time under the same mathematical model.

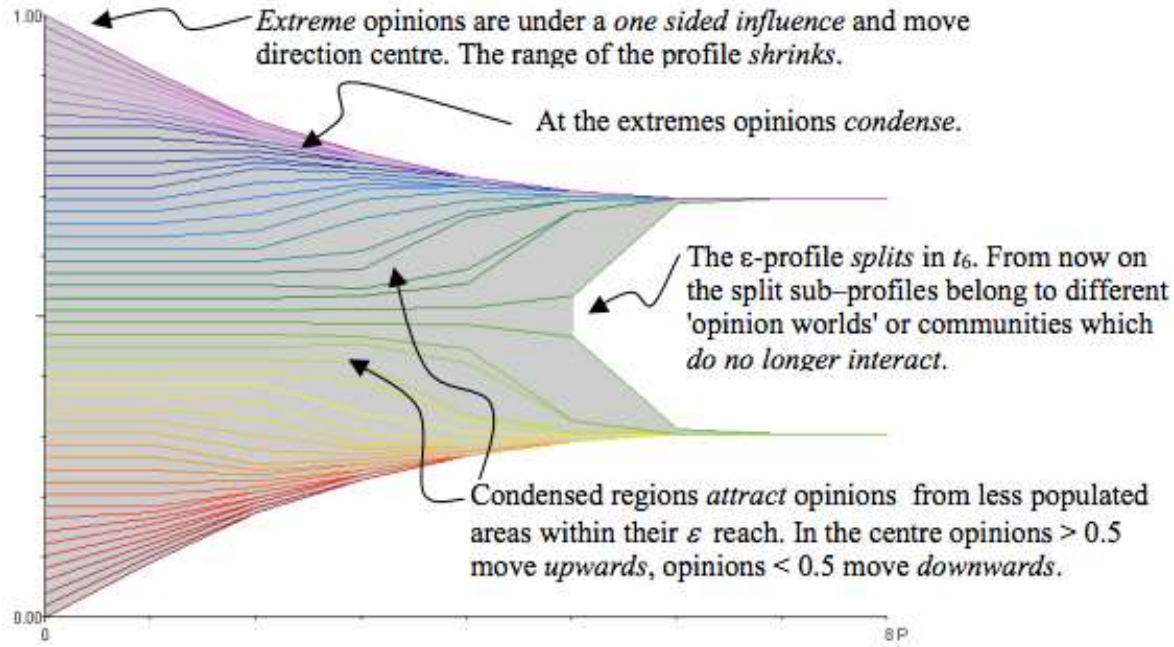


Figure 2: Displays the converging of opinions into isolated communities over time.

It can be seen that agents on the extremes of the opinion spectrum (close to 0.00 or 1.00 in the above figure) are incredibly unlikely to cross over to opposing extremes. The best that can occur is for extremely opinionated agents to become averaged with more local opinions. Thus, for an advertisement

campaign, this indicates that a 'realistic success' is will be a measure of how many agents are convinced to be within locality of the desired opinion space.

This model of scoring is similar to what was chosen by the evaluation matrix; a measurement prioritizing nearby agents, as opposed to considering the entirety of agent distribution. This method is referred to as calculating within an advertisement radius, and is in line with how consumer behaviour is perceived. It is logical to understand that a consumer's opinion would drive them to purchase products of similar nature to their opinion coordinates. As such we can assume that the number of customers is given by the following set C.

$$C = \left\{ c = (x_c, y_c) \middle| \sqrt{(x_c - x_a)^2 - (y_c - y_a)^2} \leq r \right\}$$

Equation 1: The equation guiding the determination of customers in the opinion plane.

This is simply the set of all points c, such that the magnitude of their distance away from the ad's location, (x_a, y_a), is less than or equal to the radius r. The determination of the radius can be taken as an estimation of the client company's dominion over the surrounding opinion space. The visualization of this radius within the model can be seen below in **Error! Reference source not found.**.

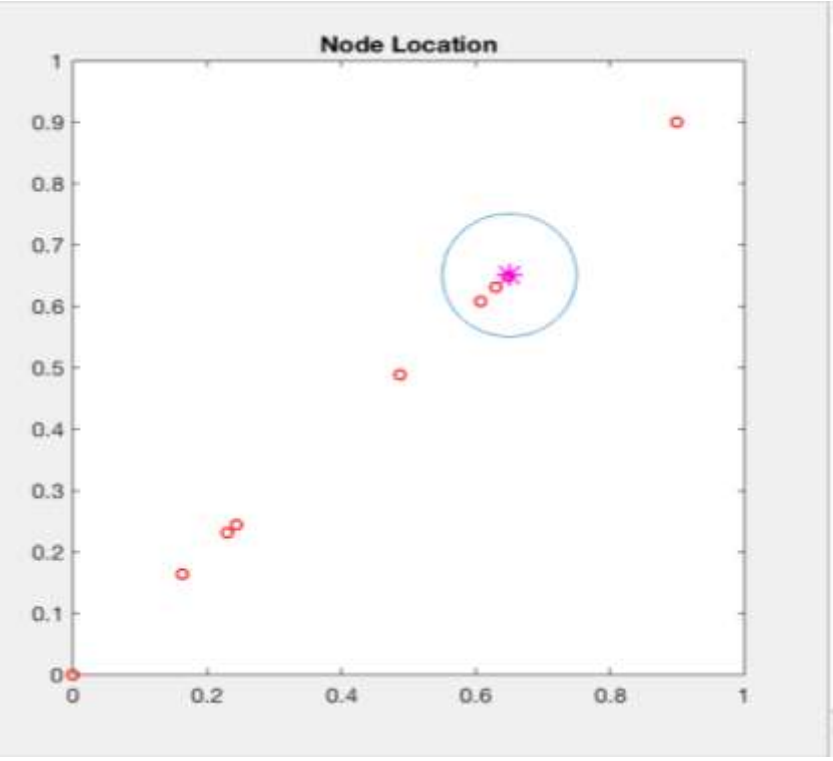


Figure 3: Displays the circle within which agents are considered to be consumers.

Thus, it is possible to take into account the presence of nearby competitors vying for similarly placed agents simply by narrowing the radius, resulting in a more rigorous scoring. This consideration of nearby agents also aligns with Al Ries and Jack Trout's book, *The 22 Immutable Laws of Marketing* [25]. In this marketing guideline, it highlights the importance of *The Law of the Mind*, by which to success in marketing is interpreted as a measurement of owning space in a consumer's mind. Thus, it is seen that modeling a company's ownership over an opinion space is an accurate representation of marketing success.

Agent Interaction

As the mathematical model continually averages the opinions of agents to define their convergence, it is crucial that this is performed in a realistic way. The solution selected by the opinion formation represents an important facet of opinion formation; modeling the *trust* that exists between individual agents. This concept of trust in the model can be thought of as a ratio by which an agent is receptive to the opinions of another agent. Agents which share a low trust value will influence each other to a lesser degree when opinions are averaged. This is similar to a report on agent interactions which defines an operationally similar 'influence' factor [26]. This is represented in the following equation for single dimension averaging:

$$x_{i_{new}} = \sum_{j} \emptyset_{ij}(x_j - x_i)$$

Equation 2: An interpretation of the Krause averaging function with an added influence factor.

In this interpretation, the value 'Ø' is representative of the trust value shared between agents 'i' and 'j', and is multiplied into the respective differences between agent locations. A similar measurement of influence or trust is included in the model created by this project. This weighted sum provides a more realistic convergence of opinions, and can be replicated through the use of data taken on the degree and quality of interaction between agents. Thus it becomes apparent that the model will prioritize not only agents that are within the radius of other agents, but also that are inherently more influential by this measure.

Agent Iteration

The final engineering specification for this model is the iterative method of the model. That is, how the model selects which agents to score as it ranks each for determination of the ideal influencer. The

method selected by the evaluation matrix was a complete iteration, which entails the model running through as many iterations as there are agents. While this may seem an inefficient use of computing power, this ensures that the space is entirely analyzed. This full analysis becomes especially important when taking into account the incomprehensible quantity of potential outcomes. Every opinion space will host a quantity of agents, each with a unique radius and an array of associated trust values. With this degree of variation in mind, it becomes simpler for the current stage of the project to simply iterate through each agent as the target agent to determine the maximum score.

Additionally, the model operates under the assumption that it will only extract arrays of data which are properly sorted. As such, to iterate the data stands out as the ideal solution according to a Cornell study [27]. Iteration becomes optimized when the data set is well organized. Rather than utilize a pattern recognition function to sort through the arrays of radii and associated trust values in an attempt to discount certain agents, it in fact is ideal to merely test out each assumption. When data sets become particularly massive, it may be ideal to switch to a method of attempting to negate the testing of certain agents, but this will be discussed in section 4.5 Recommendations.

4.3 Project Economics

The need to acquire data for the algorithm is a critical aspect for the success of this project. As such, it is vital to look into the proposed economics of the model. To start, the cost of online advertising will be explored. Most common advertisements (search based or banner ads) are priced based on their click-through rate. For example, the average ad on Google costs $2.32 per click, while the average ad on Facebook costs $1.72 per click [28]. Google charges approximately $59.18 for an advertising campaign to be attached to certain keyword searches. In addition, companies can pay more for their advertisements to be shown higher on the Google search list based on keywords. For an ad to be appear on the first page of a Google search, it costs a minimum of $50 per click [28].

Another important economic consideration is the cost of agent data. As gathered from Financial Times, information as simplistic as location, age and gender of any given user can be sold for approximately $0.0005 ($0.50 per thousand users) each [29]. If an agent is deemed “influential” by product of having a dedicated following or being a public figure, these prices can be valued upwards of $0.00075 ($0.75 per thousand) [29]. The term influential is relative to the environment in which the company wishes to utilize the purchased information. Data becomes notably more expensive as it becomes insightful. These sets include: shopping history, data regarding income, and information surrounding distinct demographics (level of education, marital status) [29]. Companies such as LeadsPlease can sell data

containing the names and mailing/email addresses of agents that suffer from specified medical conditions, such as: diabetes, cancer, and obesity, for upwards of $0.26 per agent [29]. From this information, the cost of collecting various fields of data is seen to be generally inexpensive.

It is similarly important to consider the implications of marketing analytics on a company's salience and advertising effectiveness. As gathered from the August 2015 CMO Survey, companies spent 6.7%, and are projected to spend 11.1% over the next three years (in relation to 2015) of their advertising budgets on market analytics [30]. Utilizing the CMO Survey as a source for data, a study was conducted to determine the effectiveness of these methods of analysis, and their effect on annual profit. The results concluded that the application of market analytics produced average increases of 0.39% in profits and 0.61% in marketing ROI, per applied area [30]. To elaborate, if a company applied marketing analytics to 2 of the 11 identified areas present in the survey, then that would yield a 0.78% increase in profits and 1.22% increase in ROI [30]. The 11 areas in which companies are using marketing analytics is presented in **Error! Reference source not found.** [30].

Figure 4: Categories of applied marketing analytics.

For the design model, the ROI was calculated using the equation: $ROI = \frac{Gain - Cost}{Cost} \times 100$. The gain of the model can be approximated from the built-in scoring algorithm, whereas the cost can be valued by approximating the data's worth for all 10 agents and the cost of one targeted ad. Using the calculator in [29], the data of each agent will be assumed to be $0.05, yielding a total cost of $0.50 to purchase relevant information on any given 10 agents. It is important to note that the number of agents used is strictly arbitrary, and the number 10 has no real significance. For simplicity, the social media platform will be assumed to be Facebook, where the cost per click of an advertisement is roughly $1.72 [28]. Under the further assumption that the agent who gets sent the ad is guaranteed to click on it, and said

agent is the only one to see the ad, then that equates to $1.72. Hence, the cost associated with this system would be $2.22. In larger systems, more agents would see the ad, thus increasing the cost associated with it.

After several tests of the model, it was determined that the average score of any given agent (non-optimized) was 4.6. Conversely, the average score of an optimal agent was 8, thus yielding a gain of 3.4. Since the model remains generic, the variable 'n' denotes the profit attributed to the sale of the product. The net gain can then be represented as $(3.4)n$. Thus, the ROI of the model is $ROI = \frac{(3.4)n-2.22}{2.22} \times 100$. This implies that even if the profit associated with the product is only $1, the ROI would be 53%. The reason the ROI is significantly higher than those present in real-world applications is because the model has made numerous over-simplifications, producing an ideal model. The most notable of these assumptions being the two-dimensional opinion-scape.

4.4 Assessment of Final Design Project Metrics

Upon competition of the design solution, it is imperative to assess the functionality of the model with respect to the design criteria set forth in the initial design process. The following criteria were identified within 2.2 Design Criteria and Specifications, as shown in Table 12: Contains the project metrics to assess the final design solution by.:

Table 12: Contains the project metrics to assess the final design solution by.

Criteria #	Description of Criteria
1	Flexibility: The model's ability to represent an opinion space is only limited by the computational power of the electronic system.
2	Data Security: The data used in the model is not at risk of malicious activities (i.e. data theft, hacking, etc.)
3	Efficiency: The use of the model is an efficient use of computing resources and time.
4	Opinion-Scape Dimensionality: The model is able to consider a realistic number of dimensions in its calculations.
5	Scoring: The model is able to properly evaluate an agent's performance and provide a clear and distinct 'optimal agent'.
6	Accuracy: The model's output will closely model how ad targeting will affect an opinion space.

Criterion 1: Flexibility

Flexibility is defined by the ability of the model to scale and adapt to any desired application from a set of prepared inputs. In its current state, our model can effectively receive any inputs from an excel file allowing it to generate any unique space based off a company's needs. The number of agents and their radii are directly proportional to the amount of data that the company collects on its users. Through the

file extraction and using effective size-determining variables in the code, the model's resulting graph will always properly represent the company's needs.

Criterion 2: Data Security

Data security is the ability for external parties to breach the private information that is used within the model. The excel spreadsheet that stores the agent's personal information will be encrypted to ensure the security of data. As previously discussed in,

4.1 Interface Design, files will be transferred over secure networks and stored in the cloud with encrypted access keys. Thus, the model will provide proper data security in its interactions with all stakeholders of the project.

Criterion 3: Efficiency

Efficiency is a measure of how well the model utilizes resources (including time and computing power) to produce the output of the ideal agent. The model's process currently iterates through all of the agents to determine the optimal marketing target. This process is performed with a $O(n^3)$ efficiency grade, which is the most efficient method as was determined in 4.2 Model Design. The alternative of performing an analysis on each agent's accompanying data would result in an even greater number of iterations, thus highlighting complete iteration as the most effective method.

Criterion 4: Opinion-Scape Dimensionality

The dimensionality of the opinion scape is a measure of the model's ability to represent unique stances within the opinion plane. The current model only contains two dimensions, and is thus capable of representing four different unique stances. Operating under the assumption that there are only two stances which determine an agent's loyalty to a brand is inherently wrong, and an ideal system would be able to increase the number of dimensions to better represent the complexity of human opinions. Thus, this criterion is not met to its full extent, however under the assumptions of the project, the goal to model a two-dimensional space was met.

Criterion 5: Scoring

The model was successfully able to score all agents, targeting them each individually and gauging their influence over others with a quantitative scoring method. The scoring method, discussed in 4.2 Model Design, takes a realistic approach to quantifying the increase in salience for the client company. As such this criterion is met in full.

Criterion 6: Accuracy

The ability of the model to maintain accuracy, is to a certain degree lost in its assumptions. To properly predict how an agent will react to a certain form of advertisement is dependent on the quality and nature of the data collected. This topic of data collection is outside the scope of the project. There are areas by which the model maximizes on accuracy. The method of scoring accounts for a competitive market, as only agents directly surrounding the ad's opinion space are considered successful consumers. As well, there is level of trust between the agent and the advertisement, which is representative of real world interactions based on a variety of studies mentioned in 4.2 Model Design. Thus, within the scope of the project, this model satisfies the criterion of accuracy.

4.5 Recommendations

As it stands, this project's model is built to optimize itself within the current set of assumptions. To propel the model forward, it becomes critical to strip away these assumptions and embellish the model with appropriate components that will contribute to a more realistic outcome. Knowing this, future aspirations of the model include: increasing the dimensionality of the model, accounting for a competitive market, and across the board, further research into the applications of available data to model a time-variable space.

In real-world applications, the complexity of human behaviour and decision making exceeds the limitations of a strictly two-dimensional model. Agents can have dozens of subtle viewpoints on any given topic, and those subtleties are often what lead to accepting or rejecting an idea. The increase of dimensions represents a simple expansion upon the application of Krause mathematics; to use variables beyond the X-Y plane. It is important to note that with the expansion of dimensions comes a need to acquire more data pertaining to each additional axis in order to accurately place the agents. Adding these dimensions would greatly expand the model's accuracy, as well as allow this project to model more complex spaces.

Furthermore, the present model operates under the assumption that the advertised brand has no competitors in the space. Although the scoring method attempts to optimize on this assumption by only considering local agents, a more realistic consideration would consider the effect of competitive advertisement on the convergence of agents. This too comes with the need to use data, as it is necessary to place competitors accurately on the plane.

As previously mentioned, the intricacies of human behaviour are ever-changing, and modelling them as constant values is inaccurate. Thus, the trust values between agents would be more accurately modelled as a function of time. Using empirical trust metrics, trust values can be optimally calculated by extracting more data for each given agent. Pertinent data could include: direct and indirect interactions between agents, religious and political stances, transportation patterns, and other ideologies. Another recommendation for the model is for the location of the ad itself to become a function of time. This concept of an 'ad trajectory' is recommended to incorporate an option for the ad to move along a pre-destined path. This method of marketing provides the opportunity for long-term marketing campaigns, with the intent to build more consistent brand awareness as opposed to the one-time conversion of sales or website views. For simplicity, the path should be implemented as a linear function of time to determine whether or not that truly is more effective than a static model. As more information and

experience is gathered, more complex functions of time can be implemented to calculate the optimal way of advertising to any given set of agents.

4.6 Conclusions

The outcome of this project produces a system that can identify the top influencer within a given social system. This comes with it the opportunity to drastically increase the effectiveness of advertisements in online marketing. The designed model is adaptable to any two-dimensional opinion plane, as it can take any scale of user data an input and iterate through all data sets to produce the optimal agent. As it is processed through MATLAB, the model comes with detailed visual simulations, which provide insight regarding the path of convergence of each agent.

Using basic information on agents the model is capable of iterating through and scoring agents based on their ability to influence others. The algorithm brings a new element to online marketing as it introduces a new method for ad targeting. This provides the ability to connect companies with ideal consumers, increasing the saliency of a company and ideally supporting the consumer with a desired product or ideology. The design permits a decrease in advertising costs for companies especially beneficial for those with low budgets, such as start-ups. This in turns helps smaller companies compete with large scale corporations who typically can reach more agents as a by-product of their increased budgets.

The model was a success in the sense that it was adaptable and able to identify the top influencer of a given group, whilst maintaining the safety and security of agent information. However, due to the surplus of oversimplifications and assumptions, the model's accuracy was compromised. Thus, to be applied in real-world situations, it is highly recommended that the complexity of the model be raised significantly before use.

References

[1] M. R. Johannessen, Ø. Sæbø and L. S. Flak, "Social media as public sphere: a stakeholder perspective," Emerald Group Publishing Limited, Bradford, 2016.

[2] C.-N. Ziegler and G. Lausen, "Propagation Models for Trust and Distrust in Social Networks," 2005. [Online]. Available: http://citeseerx.ist.psu.edu/viewdoc/download;jsessionid=6CCEA41546C80A1E27FAB200915FE7C1?doi=10.1.1.85.8407&rep=rep1&type=pdf. [Accessed 2 November 2017].

[3] NORC, "The General Social Survey," University of Chicago, 2016. [Online]. Available: http://gss.norc.org/About-The-GSS. [Accessed 2 November 2017].

[4] E. L. Glaeser, D. I. Laibson, J. A. Scheinkman and C. L. Soutter, "Measuring Trust," 2000. [Online]. Available: https://scholar.harvard.edu/files/laibson/files/measuring_trust.pdf. [Accessed 2 November 2017].

[5] C. Pumphrey, "How do advertisers show me custom ads?," 2012.

[6] MassLive Marketing Staff, "How Does Audience Targetting Work?," MassLive Media, 2017.

[7] E. Asano, "How Much Time Do People Spend on Social Media," 4 January 2017. [Online]. Available: https://www.socialmediatoday.com/marketing/how-much-time-do-people-spend-social-media-infographic.

[8] R. Bell, "A beginner's guide to Big O Notation," 23 June 2009. [Online]. Available: https://rob-bell.net/2009/06/a-beginners-guide-to-big-o-notation/. [Accessed 5 November 2017].

[9] M. Anderson, "Technology Device Ownership: 2015," 2015.

[10] S. Ganguly, "E-WASTE, GENERATION, DISPOSAL AND ITS ASSOCIATED ENVIRONMENTAL IMPACT: A REVIEW," 2014. [Online]. Available: http://www.indjsrt.com/administrator/modules/category/upload/2-INDJSRT201405-05.pdf. [Accessed 1 December 2017].

[11] J. Jercinovic, "The Ethics of Using AI In Advertising," 2017.

[12] Z. A. Devraj, "Epic Rise of Digital Marketing in Industry- Are You Ready to Invest?," 2017.

[13] Pew Research Center, "State of the News Media 2016," 15 June 2016. [Online]. Available: https://assets.pewresearch.org/wp-content/uploads/sites/13/2016/06/30143308/state-of-the-news-media-report-2016-final.pdf. [Accessed 1 December 2017].

[14] QuickBooks, "Is My Data Safe In The Cloud?," 2016. [Online]. Available: https://quickbooks.intuit.com/r/tech-review/is-my-data-safe-in-the-cloud/. [Accessed 30 November 2017].

[15] M. Wall, "Can we trust cloud providers to keep our data safe?," BBC News, 2016.

[16] M. Sweeney, "When Can You Use PII for Online Advertising and Marketing?," February 2017. [Online]. Available: https://piwik.pro/blog/pii-online-advertising-marketing/. [Accessed 1 December 2017].

[17] S. Vizard, "New online ad laws 'put future of the web in danger'," 2017.

[18] S. Pate, "Top 10 Encryption Benefits," 30 April 2013. [Online]. Available: http://www.infosecisland.com/blogview/23115-Top-10-Encryption-Benefits.html. [Accessed 1 December 2017].

[19] J. Lock, "Canadian Privacy Laws and the Canadian Cloud: A Primer for Canadian Businesses," 15 September 2015. [Online]. Available: https://www.servercloudcanada.com/2015/09/canadian-privacy-laws-canadian-cloud-primer-canadian-businesses/. [Accessed 1 December 2017].

[20] Bokardo, "Principles of User Interface Design," [Online]. Available: http://bokardo.com/principles-of-user-interface-design/. [Accessed 1 December 2017].

[21] A. Watson, "Long Term Advertising Strategy in a Short Term World," 7 July 2016. [Online]. Available: https://www.adclarity.com/2016/07/long-term-advertising-strategy-short-term-world/. [Accessed 1 December 2017].

[22] S. M. Jafari, "Impact of trust and perceived content of advertisement on intention to accept mobile advertisement," 2016.

[23] E. Shultz, "What Kinds of Advertising Do Consumers Trust the Most?," 25 September 2013. [Online]. Available: http://adage.com/article/news/kinds-advertising-consumers-trust/244360/. [Accessed 1 December 2017].

[24] J. Ouellette, "What Does It Take to Change a Mind? A Phase Transition," 13 February 2015. [Online]. Available: https://blogs.scientificamerican.com/cocktail-party-physics/what-does-it-take-to-change-a-mind-a-phase-transition/. [Accessed 1 December 2017].

[25] A. Ries and J. Trout, The 22 Immutable Laws of Marketing, New York, NY: HarperBusiness, 1993.

[26] A. Aydogdu, M. Caponigro and S. McQuade, "Interaction Network, State Space, and Control in Social Dynamics," 2016.

[27] Cornell University, "Recursion v.s. Iteration. Which is Better?," [Online]. Available: http://www.cs.cornell.edu/info/courses/spring-98/cs211/lecturenotes/07-recursion.pdf. [Accessed 1 December 2017].

[28] D. Shewan, "The Comprehensive Guide to Online Advertising Costs," 2017.

[29] E. Steel, "Companies scramble for consumer data," New York, 2013.

[30] M. Ariker, A. Diaz, C. Moorman and M. Westover, "Quantifying the Impact of Marketing Analytics," Harvard Business Review, 2015.

[31] "How Much Does Social Media Marketing Cost?," 2013.

[32] D. G. Mikulski, "Cooperative Trust Games," in *Game Theory Relaunched*, Intech, 2013.

[33] S. P. Marsh, "Formalising Trust as a Computational Concept," April 1994. [Online]. Available: http://www.cs.stir.ac.uk/research/publications/techreps/pdf/TR133.pdf. [Accessed 2 November 2017].

Appendix A: Quality Function Deployment

Table 13: Holds the weightings of each customer requirement from the perspective of stakeholders.

Stakeholders:	Ad Agencies	High Salience Companies	Low Salience Companies	Owners of Data	Consumers
Customer Requirements					
Scalable	18	19	16	~~	~~
Adaptable	18	17	16	~~	~~
Accurate	20	20	18	13	~~
Cost-Effective	13	17	17	~~	~~
Optimized Output	16	19	18	~~	~~
Increase to Saliency	15	9	19	10	~~
Security of Information	11	16	15	20	20
Ethical Data Use	12	6	9	20	20
Non-Intrusive	7	6	9	16	19

Table 14: Displays the results of the QFD for four of seven engineering specifications.

Stakeholders:	Weight	Method of Data Input	Scoring Algorithm	Ergonomic Features
Customer Requirements				
Scalable	17.7	9	6	6
Adaptable	17	6	~~	3
Accurate	17.8	9	9	~~
Cost-Effective	15.7	6	~~	6
Optimized Output	17.7	~~	9	~~
Increase to Saliency	13.3	~~	6	~~
Security of Information	16.4	9	9	~~
Ethical Data Use	13.4	6	6	~~
Non-Intrusive	11.4	~~	~~	~~

Table 15: Displays the results of the QFD for four of seven engineering specifications.

Stakeholders:	Weight	Agent Interaction	Ad Interaction	Agent Iteration
Customer Requirements				
Scalable	17.7	~~	~~	6
Adaptable	17	~~	~~	~~
Accurate	17.8	9	9	3
Cost-Effective	15.7	~~	9	6
Optimized Output	17.7	9	9	9
Increase to Saliency	13.3	9	9	~~
Security of Information	16.4	6	6	~~
Ethical Data Use	13.4	~~	~~	
Non-Intrusive	11.4	~~	9	~~

Appendix B: Evaluation Matrix

Table 16: Displays the evaluation matrix for considered design solutions

	Engineering Specification	**Method of Data Input**			**Ad Interaction**		
	Design Solution	File Extract	Manual Input	Interpolate	Limited Range	Ad Trust Values	Ad Trajectory
Criteria	Weight						
Scalable	13.7	4	1	3	2	2	3
Adaptable	13.7	4	4	4	2	3	2
Accurate	9	4	3	1	1	4	4
Cost-Effective	15.6	3	1	3	3	2	1
Optimized Output	14.4	3	3	1	2	4	4
Increase to Saliency	14.9	3	3	2	2	3	3
Security of Information	11.6	2	3	4	3	2	2
Ethical Data Use	13.5	2	2	4	2	2	1
Non-Intrusive	17.8	1	1	1	4	2	1
Totals		348	279	314	302	324	277

Table 17: Displays the evaluation matrix for considered design solutions

	Engineering Specification	Scoring Algorithm			Agent Interaction	
	Design Solution	Ad Radius	Convergence of Agents	Net Change	Agent Trust Values	Equal Consideration
Criteria	Weight					
Scalable	13.7	3	2	4	2	4
Adaptable	13.7	4	2	3	3	2
Accurate	9	4	3	2	4	1
Cost-Effective	15.6	4	3	2	2	2
Optimized Output	14.4	4	2	2	4	2
Increase to Saliency	14.9	3	4	2	4	1
Security of Information	11.6	2	2	2	1	4
Ethical Data Use	13.5	3	3	2	2	4
Non-Intrusive	17.8	4	3	2	2	2
Totals		432	334	290	327	302

Table 18: Displays the evaluation matrix for considered design solutions

	Engineering Specification	Agent Iteration		Ergonomic Features	
	Design Solution	Complete Iterations	Sorted Iteration	Ad Location Input	Ad Trust Slider
Criteria	Weight				
Scalable	13.7	2	3	4	3
Adaptable	13.7	4	3	4	3
Accurate	9	4	4	3	4
Cost-Effective	15.6	3	2	2	2
Optimized Output	14.4	4	3	3	3
Increase to Saliency	14.9	4	3	4	3
Security of Information	11.6	2	1	2	1
Ethical Data Use	13.5	2	1	2	1
Non-Intrusive	17.8	1	2	2	2
Totals		350	298	356	298

Appendix C: Evaluation Matrix Rubric

Table 19: Evaluation Matrix rubric.

Criteria	1 (Poor)	2 (Average)	3 (Good)	4 (Excellent)
Scalable	Model is unable to scale to any degree, and is only capable of portraying a space with a specific number of agents.	Model is able to scale from a set quantity of agents only within a limited range of approximately 5-20 agents.	Model is able to scale from a set quantity of agents to within a limited range of agents. The limited range negates the portrayal of certain opinion scenarios.	The model is able to scale entirely based on the required opinion space. There is no limit to the quantity of agents portrayed aside from the data perceived and physical limitations of computation.
Adaptable	The model can only be applied to incredibly specific two dimensional scenarios, discounting a wide variety of uses.	The model is able to be applied to a general category of two-dimensional opinion scenarios, with a variety of opinion scenarios that it cannot generate.	The model is able to be applied to most two dimensional scenarios, with a select few that it cannot generate.	The model can be applied to any two-dimensional opinion scenario, and is only limited by the data to which it is given.
Accurate	The model's results are not accurate for the use of identifying the top influencer in the opinion plane. The algorithm's result contradicts what is modelled by the Krause mathematics.	The model is able to identify a top influencer that appears to be in the top tier of possible influencers. There is no consensus between Krause mathematics and the algorithm as to which agent is the top influencer.	The model is able to identify a top influencer which appears to increase a company's saliency or 'opinion profile'. Krause mathematics employed in the model align with the predictions of the algorithm used to determine the top influencer.	The model represents the top influencers in an opinion plane to an incredibly accurate extent. What is modelled by Krause mathematics is supported by the algorithm's prediction of top influencers. Notable increases to a company's saliency occur with the selection of the identified influencers.
Cost-Effective	Model requires laborious input of data from the perspective of manpower. Algorithms are computationally inefficient, resulting in an ineffective use of resources.	Model requires some intervention alongside autonomous features. Algorithms are not optimized for efficiency but are not wasteful of computing resources.	Model is mostly autonomous aside from user interface components. Algorithms are efficient and not wasteful of computing resources.	Model is entirely autonomous aside from user interface components. Algorithms are incredibly efficient with an excellent use of computing resources.
Optimized Output	The results and suggestions of top influencers are portrayed in ways that render them unusable. Further computation is required to receive	The proposed suggestion of top influencers is somewhat interpretable without further computation. Solution does not	Suggestions of top influencers is able to provide value to the advertising campaign's direction. Information received is usable and employable. Solution	Suggestions of top influencers provides immediate value to the advertising campaign's direction. Information provides insight into the opinion plane, allowing for greater strategic

	useful information from the model.	optimize on quantity of ads deployed.	optimizes mildly on quantity of ads deployed.	direction. Solutions optimize greatly on quantity of ads deployed.
Increase to Saliency	Companies receive negative or no benefit to their saliency by following the suggestions of the model.	Companies will receive minimal benefit to their saliency through following the suggestions of the model.	Companies will receive a mild benefit short term to their saliency through following the suggestions of the model.	Companies will receive a significant and permanent benefit to their saliency through following the suggestions of the model.
Security of Information	As information is passed to the model from the data owner it is placed in significant jeopardy of being stolen, hacked, or altered.	Information being received by the model is placed in a position of mild risk from third parties.	Information being used by the model is under minimal risk of being stolen, hacked, or altered by most parties.	Information being passed to the model is under no risk of being altered, stolen, or hacked for malicious purposes by any parties.
Ethical Data Use	The data used by the model crosses ethical boundaries, and is collected in a manner that it is used in unethical marketing practices.	The data used by the model is unethical to a minimal extent. Slight risk of unethical marketing practices.	Data use in the model is for the most part ethical, and is not used to pursue unethical marketing practices.	Data use in the model is entirely ethical. There is no risk that data will be used to pursue unethical marketing practices.
Non-Intrusive Advertising	Advertisements appear invasive to individuals within the social system.	Advertisements are minimally invasive for individuals in the social system.	Advertisements are relevant to individuals in the social system, and do not appear as though invasive.	Advertisements are integrated well in the social system. Agents do not perceive the advertisement as invasive.