An adaptive temporal-causal network model to analyse extinction of communication over time

Lucas Johannes José Fijen\textsuperscript{a}, Julio Joaquín López González\textsuperscript{b}, Jan Treur\textsuperscript{b}

\textsuperscript{a} University of Amsterdam, Computational Science, the Netherlands
\textsuperscript{b} Vrije Universiteit Amsterdam, Social AI Group, the Netherlands

Received 17 July 2020; accepted 26 August 2020
Available online 19 March 2021

Abstract

The persistence of information communicated between humans is difficult to measure as it is affected by many features. This paper presents an approach to computationally model the cognitive processes of information sharing to describe persistence or extinction of communication in Twitter over time. The adaptive mental network model explains, for example, how an individual can experience information overflow on a topic, and how this affects the sharing of information. Parameter tuning by Simulated Annealing is used to identify characteristics of the network model that fit to empirical data from Twitter. The data collected is related to the independentism in Catalonia, Spain, which is considered a global issue with repercussion in Europe.

© 2021 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

Keywords: Extinction of communication; Network modeling; Adaptive network; Social simulation

1. Introduction

Nowadays, everyone is communicating using new technologies. The Internet has introduced a full new level of information transmission. Every single person, even elderly people, can have a video chat with their family members or post a comment about any kind of topic. Especially, social media sites have dramatically changed the ways people connect with each other (Liu & Ying, 2012). Particularly, social networks now make emphasis on how the user-content is created to be consumed by their connections (Ellison & Boyd, 2019).

Before social media, the spread of information was limited by the infrastructure of traditional media. When something important was happening in the other part of the world, it was hard to be informed or prepared about it. Human social interaction and the Internet have removed this limitation, and now there is a vast amount of information flowing in all directions about a broad range of topics. Now with the reach of Facebook or Twitter, these social networks have become some pseudo-news agency, an intersection between news and social networks, where people go to know what is happening in the world (Mitchell, Rosenstiel, & Christian, 2012).

The process of sharing information and using it to reflect ideas is something that is intimately related to social networks, especially Twitter. Users on Twitter receive information from a large part of the society, from news agencies to relatives and friends. All of this can have an influence on the decision of retweeting (Hill, Lo, Vavreck, & Zaller, 2013).
This paper aims to develop a computational understanding of how news events all over the world persist over time. This will provide a deeper understanding of the cognitive processes that drive a person to share information about an event and how the persistence can change depending on the person. To do so and based on the Network-Oriented Modelling for Adaptive Networks approach described in (Treur, 2020), an adaptive network was designed to model the willingness of a person to share information related to an event. The model represents how the information-sharing behaviour can be accelerated and decelerated, given the closeness a person has to a topic and the popularity of such a topic, but also the extent of information overload, exhaustion or being fed-up of the person. Example simulations illustrate how these factors affect the extent of sharing tweets and how the persistence of news changes between different types of persons. The model has been analysed and validated by comparing it to empirical information from Twitter.

The paper is organised as follows. First, Section 2 discusses the literature on sharing behaviour in some depth and the study is situated, giving a focus target for the study. In Section 3, the adaptive temporal-causal network model is introduced and explained. Section 4 explains the data sets used in the study, followed by the experimentation section, where the tuning experiments using the data set are explained. In Section 5 it is discussed how verification of the model by mathematical analysis of stationary points was performed. Finally, in the discussion Section 6 the main findings for the model are shown.

2. Drivers for sharing behaviour

In order to develop a behaviour model that can explain how information persists in the world, a basic understanding of the underlying concepts for sharing behaviour must be established. In this case, there is a difference between offline processes and online processes. The latter ones focus on Twitter sharing behaviour as it is the most important sharing behaviour that the network will account for. Therefore, the most basic ways of sharing are explained. Different factors can affect the behaviour such as the source from where an event is coming, the popularity of a tweet, here the engagement of a tweet is elaborated more in-depth. Offline behaviour processes are also crucial for describing proper behaviour. There is usually a balance between online interactions and offline interactions, and it will be explained more in-depth on how this relationship between these two factors is built. Finally, there is a last important factor called overexposure to information or information overload (Eppler & Mengis, 2004). Previous studies have determined that too much information can damage the engagement and sharing behaviour over time by reducing the number of interactions and thus creating a negative behaviour toward new events related to the same idea.

When an event happens, the emotional response of those affected by it is deeply connected with their sharing behaviour. Sharing behaviour in Twitter is considered everything from tweeting about a topic, retweet about it, or to mention other users that have connections with the topic. However, online sharing behaviour depends on more factors than only the emotional response (Kim & Yoo, 2012). Important factors are the content of the information (information and multimedia), the user and the network characteristics (if the source of information is an influential person/news agency) and the current popularity of a specific event (how much people have been talking about an event in the last week). The different factors that can affect the sharing behaviour are explained in depth in the next section.

Several studies try to understand what makes someone more engaged with a topic, particularly the underlying motivation for retweeting among twitter users (Boyd, Golder, & Lotan, 2010; Lee, Kim, & Kim, 2015; Recuero, Araujo, & Zago, 2011; Suh, Hong, Pirolli, & Chi, 2010). Suh et al. (2010) collected 74,000 tweets to understand which elements are more critical when deciding to retweet or not. He identified tweets with URL and hashtag, tweets from influential users (a large number of followers and followers) and users with a long story of using Twitter were more inclined to be shared. Boyd et al. (2010) focused on the problem from a psychological perspective. He identifies the action of retweeting as a way of starting a conversation and stay engaged with other users. Other researchers identified factors such as information sharing or responsiveness. Furthermore, Lee et al. (2015) consider that retweeting is motivated by prosocial reasons, especially those that have an altruistic motivation. Recuero et al. (2011) also had a similar view where the value the users perceived from their retweet was beneficial for the social network as a whole, and this can be seen as altruistic behaviour.

As seen from the review above, studies have provided different responses to the question of why Twitter users decide to share a message. The truth is that the sharing process is complex and can be driven by multiple factors. It is the combination of these factors that makes a user more engage in a topic. The engagement of an event is defined as how motivated people are to share their opinion about a specific event. The closeness to the event, the amount of offline interaction the user has in their circle, and the popularity the event has online in addition to all the factors mentioned above trigger the engagement. The current proposal also adds offline behaviour to the formula.

The term “information overload” is usually used to describe the act of receiving too much information. This concept has been applied to a variety of situations, from research to shopping, and they study how the performance of an individual is affected by the amount of information he or she is exposed to. The more information a user receives, the more performance he or she will have but only up to a certain point. After a tipping point, the performance of an individual will rapidly decline when more information is received (Chewning & Harrell, 1990). Infor-
information overload then confuses the user, thereby reducing their abilities to a small set of priorities. This construct is also present in social environments. However, it has slightly different effect, by overloading the cognitive abilities to process information of the social media users it indirectly affects the social media contagion effect (Gomez-Rodriguez, Gummadi, & Schölkopf, 2014). According to Manuel, when a user is overloaded, they prioritize tweets from a small number of sources and their interactions are mostly related with this small group. In contrast, when a user is not overloaded, they tend to increase the number of interactions and they increase the information transferred.

This information overload factor has essential repercussions in the information propagation and persistence of ideas over time. There is a large number of studies that focused their research on information sharing in social media (Goel, Watts, & Goldstein, 2012; Gomez-Rodriguez, Leskovec, & Krause, 2012; Schoenmaker, Treur, & Vetter, 2018), but most of them ignore the information overload and how it can accelerate or decelerate the information contagion effect. These studies consider that the spread of information is motivated by only them, ignoring that there can be lots of background information simultaneously being speeded through the network. In this paper, the background information affecting users is also taken into account in the model.

Catalonia is a region situated in the north of Spain that is currently living a process of becoming independent. Catalan independence started in 1923, but it had not become global until the 27 of October of 2017 when the Catalan government issued a Catalan declaration of independence. This declaration did not receive recognition from the international community. Since then, there have been several problems and events between the Spanish government and Catalonia to try to solve the issue in a pacific way (“Independencia de Catalunya | EL MUNDO.”, 2019; Lewis & Perez Maestro, 2019).

Catalanian independence is an important event as it can have repercussions at a worldwide level. Many countries are expectant to see how the big powers of the world react to the movement as if the independence movement becomes a success, many other regions in the world will try to become independent, starting a “Snowball effect” that can affect many countries. Moreover, there are already some new movements that gained strength after the Catalan referendum, places such as Scotland (Bonnett, 2019) and Hong Kong (CNN, 2019).

From a research point of view, Catalan independence is an interesting event as it allows to observe how an ongoing event is being covered and received all around the world. There has been a large number of events since 2017 (Reuters, 2019), and the process is still ongoing. Remarkably, in the few days after the process, there were a large number of tweets from different countries rejecting or supporting the events, and, as of December 2019, there is still an amount of around 2000 English-only tweets per week talking about the situation.

It also allows observing how the information sharing works outside of the limits of the countries. Spain’s Official language and national language is Spanish, a language that while even being popular is not the main language for communication in Europe. Thereby, it is interesting to observe how foreign (English speakers) countries respond to a situation and to what extent they talk on social networks about something that is only part of Spanish internal affairs.

3. The adaptive temporal-causal network model

In this section a computational model is introduced to analyse the tweeting behaviour with respect to the persistence of news tweets on the Catalanian Independence. The used model is a temporal-causal network model with a single adaptive reification level (Treur, 2016, 2020). A temporal-causal network model is constructed out of multiple states ($X_1, \ldots, X_m$). Some of these states are connected through a directed connection, indicated by an arrow in Fig. 1, and described in matrix format in Box 1. Each connection $X \rightarrow Y$ represents a causal relation from the source state $X$ to the receiving state $Y$. Throughout a simulation, the value $Y(t)$ of any state $Y$ is updated using its current value $Y(t)$, and the values $X_1(t), \ldots, X_k(t)$ of the states $X_1, \ldots, X_k$ from which $Y$ has incoming connections. For these incoming connections a weight factor $\omega_{X_i,Y}$ describing the significance of the connection, is applied as a modifier multiplied to $X_i(t)$ obtaining causal impact $\omega_{X_i,Y}X_i(t)$ from $X_i$ on $Y$. The resulting causal impact values are then combined using a combination function $c_Y(\cdots)$ which aggregates and values the relations between the different incoming impacts. To incorporate the resulting aggregated value, a speed factor $\eta_Y$ is taken into account which influences the pace at which states are updated with new information. In the end the obtained value is added to the current state value $Y(t)$. The resulting difference equation is as follows (Treur, 2016, Ch. 2):

$$Y(t + \Delta t) = Y(t) + \eta_Y [c_Y(\omega_{X_1,Y}X_1(t), \ldots, \omega_{X_k,Y}X_k(t)) - Y(t)] \Delta t$$

(1)

When modelling the persistence of communicated information, two entities are considered: an original tweet containing information on a news event, and the agent that reads the tweet, and acts accordingly. The considered behaviour concerns in how far there is persistence of a news tweet after reading it. In the introduced adaptive temporal-causal network model depicted in Fig. 1, three input states are defined for the news tweet: The Source State (SS) $X_1$, which contains information on the source that sent the tweet, an Event Popularity (EP) state $X_2$, which defines the current popularity of the event that is described in the tweet and an Event Importance (EI) state $X_3$, which defines the magnitude of the event for the world.

Those three states have causal connections to the internal mental states of the agent that reads the tweet. The
Source State is connected to a state $X_4$ of the agent that describes the Relation with the source (RS). Through this state it is possible to define the distance between the source and the agent. The Event Importance state is connected to a Topic Engagement (TE) state $X_5$, which describes the magnitude of engagement into the topic of the agent. The Relation Source, the Event Popularity and the Topic Engagement states are connected to a state that describes
the Personal engagement (PE) of the agent regarding the current topic. This state is directly connected to a state Tweet Intensity (TI) that defines the extent to which a person is going to tweet about the news related event $X_9$. Two main processes that influence the current personal engagement besides reading a tweet: Offline Interactions (OI) $X_7$, and Fed-up factor (FF) $X_8$.

The Offline Interactions have a bidirectional positive effect on the Personal engagement: having more offline interactions on a subject will cause one’s mind to think about the topic more, and vice versa. On the other hand, there is the Information Overload, which is caused by thinking a lot about a specific topic and having had a lot of offline interactions on the subject. For example, if the agent has been thinking a lot about a topic, but also experiences a significant information overload on the topic, this will more strongly influence the agents information overload, than when the agent reads the first article on the subject, and hasn’t yet thought about the topic before.

As a result, the connection weights of the connection from the Personal engagement state towards the Information Overload state and the connection towards the Offline interaction are both defined by the adaptation states $X_{10}$ and $X_{11}$ (also indicated by $W_{PE,OI}$ and $W_{PE,FF}$) at the reification level for connection weights $\omega_{PE,OI}$ and $\omega_{PE,FF}$. The behaviour of these so-called self-model states or reification states (Treur, 2020) for connection weights is defined by a Hebbian combination function, which learns to get a higher connection weight when a lot of interactions occur. The Hebbian function thus rises in value when the two connected states have a high collective firing strength, but can also decrease the value depending on the persistence factor $\mu$. The Hebbian combination function is described as follows:

$$\text{hebh}_\mu(V_1, V_2, W) = V_1 V_2 (1 - W) + \mu W$$ (2)

The parameter $\mu$ defines the persistence factor of the Hebbian learning, which influences the pace of the flexibility at which the Hebbian function adapts to new situations. For example, when $\mu = 1$, the learning effects will stay for 100% the same, which makes it impossible to adapt downwards to new situations, and when $\mu = 0.95$, per time unit 5% of what was learnt will be lost.

The Relation with source and Topic engagement states both use an id combination function which simply outputs the single incoming impact value. For the Personal engagement, the Offline interactions, the Information overload and the Tweet intensity state, a adjusted version of an alogistic function was used. The minimum value for the logistic combination function alogistic was set to a value of 0.001, which is assumed to be a basic minimal activation level of the state in rest. The formula for the logistic combination function used is as follows:

$$\text{alogistic } \sigma_x(V_1, \ldots, V_n) = \min(0.001, \frac{1}{1 + e^{-\sigma(V_1 + \ldots + V_n)}}) = \frac{1}{1 + e^{-\sigma(V_1 + \ldots + V_n)}} (1 + e^{-\sigma(V_1 + \ldots + V_n)})$$ (3)

The network structure characteristics of the model are based on the assumptions and intuition of how humans perceive the diffusion of information and how information overload makes a person fed-up. These network characteristics were specified in the so-called role matrices $m_b, m_cw, m_cfw$, and $m_cfp$ (for base connectivity, connection weights, combination function weights and combination function parameters, respectively) shown in Box 1. For example, in role matrix $m_b$ on each row for the indicated state the other states are listed with incoming connections to that state. In role matrix $m_cw$ the connection weights for these incoming connections are shown, where for adaptive connections, the reification state representing the adaptive value is indicated: the red cells with $X_{10}$ and $X_{11}$ indications. In role matrix $m_cfw$ it is indicated which combination functions are used for each of the states, and in role matrix $m_cfp$ the values of their parameters. These role matrices are used by a dedicated refined network modelling environment that has been developed in Matlab; see (Treur, 2020), Ch. 9. This environment was used for all simulations shown in this paper.

4. Validation using parameter tuning

The main objective of the experiments is to try to approximate the actual behaviour someone would expect from individuals by the model. Network characteristics are assumed to be dependent on the person and the situation. Therefore, the following characteristics were considered to be adjusted: the speed factors $\eta$ and initial values of states $X_7, X_8, X_{10}$ and $X_{11}$, the persistence factors $\mu$ of state $X_{10}$ and $X_{11}$, Simulated annealing was used for parameter search. Simulated annealing uses the root-mean-square error (RMSE) as an error function to measure the remaining deviation. It offers an effective method to find good estimations of parameters for models with a considerable number of parameters, and even if the result of the algorithm is not the global minimal one, it is guaranteed to be a good answer.

Moreover, simulated annealing is less prone to get stuck into local optima in comparison with other optimisation algorithms. The parameter value ranges for the speed factors $\eta$ and for the initial values were set between 0 and 1, and for the persistence factors $\mu$ between 0.8 and 1.

4.1. Dataset

The collected dataset consists of about 5500 tweets collected over two weeks during November – December 2019. The purpose of the data analysis is to understand which features are associated with the sharing behaviour and to study how other features that affect users at a personal level such as the engagement or the relationship with a topic have a similar effect on the sharing process. The Twitter API search endpoint was used to retrieve all the tweets

$^1$ https://developer.twitter.com/en/docs/tweets/search/api-reference/get-search-tweets
related to the data required, in this case, Catalonia independence. More precisely, five different queries were used to retrieve the tweets with words related to the independence movement, such as “Catalonia Independence” or “#FreeCatalonia”. The dataset was pre-processed to filter out every non-English written tweet. This process also removed the non-valid characters and the emojis. The result of this process yielded a set of 5474 original English written tweets with the features considered important for the research.

A set of important features among others were extracted from the filtered dataset as shown in Table 1. The first set of features (Original Author, Is Source, User Mentions and Hashtags) are related to the tweet while the second set are the features engineered for the present research.

The first four features were extracted directly using the twitter API. The next three features were derived from the tweet data and other calls to the Twitter API. The sentiment of the tweet is retrieved using textblob.\(^2\) The sentiment of tweets is considered an essential factor when sharing information. Moreover, tweets related to sad events or with a negative sentiment foment a particular behaviour among individuals (Tighe, Goldsmith, Gravenstein, Bernard, & Fillingim, 2015).

The “Relationship with the Source” and “Topic Engagement” are two user-independent factors that require individual analysis. These features are tricky to acquire due to the limit that Twitter has on the number of API calls. To compute the value of this feature several calls are required; it is necessary to get a list of the last tweet, verify if the retweeted user is considered a “source” and if the user who retweeted is following the source user.

A source is considered to be a verified Twitter user. This means that they are users popular enough for Twitter to grant them a verified account. In this research, it is assumed that verified accounts are more willing to share real information because of their popularity and their influencing position on Twitter. Consequently, to get the feature “Relationship with Source”, each tweet in the dataset is processed to observe if the posting user is following the retweeted user or the twitter mentions. Thereby, if a user follows their source of information, it is considered that he is close to the source. This feature usually has a value between 0 and 1. However, when a user is mentioning different accounts but only follow some of them, it will have a value between those numbers.

“Topic Engagement” is another essential feature used for this paper. The sharing behaviour is directly influenced by how much information is already flowing as it will motivate the individual to become more actively engage with it. To analyse this feature, it is necessary to know the number of tweets related to an event out of the last tweets from a specific user. To limit the amount of calls to the Twitter API, only the last 50 tweets from the user are taken into consideration. Averaging the number of tweets related to the topic with the total amount of tweets, 50 in this case, return a value between 0 and 1 that will approximate how engaged a user is with the topic.

The combination of the seven features, and mainly the last two, allows determining different types of individuals to be characterised by specific settings and implied patterns of the model. Each type will have a different sharing and persistence behaviour pattern, which, depending on the population, will have different effects. The analysis of how the behaviour depends on different types of individuals is covered more in-depth in the experiment section.

In addition to the features retrieved, it is also essential to determine how new stimuli are received over time. A model with constant periodic stimuli is considered to not be very representative of how real events happen in the world. As a result, to avoid using such a static function, a generator function based on the actual dataset is created. Fig. 2 illustrates a time interpolation where for each relevant tweet from a news source, the model is active for 40 time points receiving new stimuli. Note that in all figures with graphs for simulation outcomes, time is on the horizontal axis and activation levels of states are at the vertical axis.

Therefore, to obtain more valuable data, the experiments described in the experiment section use the described generator function to have an approximation of how real data behave as an input.

### 4.2. Tuning experiments

The following experiments describe how individuals with different characteristics react to information overload. The first experiment, illustrated in Fig. 3, describes a person that is quickly experiencing information overload.

When the information overload is high, the individual is immune to new stimuli, even with new events; there is no sign of social interactions related to the topic. However, after some time without having interactions related to the topic, the information overload decreases. As soon as the information overload has decreased enough, the individual starts responding again to new stimuli starting again a similar process where the individual builds up social interaction until the fed-up factor is high enough to stop interacting.

By changing the speed factor of the information overload to a lower value of 0.001, the behaviour of the individual is quite different. As a result, the new individual has different characteristics, it is less prone to information overload and requires much more offline interactions and new stimuli from twitter to build up information overload. The result of this experiment is shown in Fig. 4. One can see that after the two weeks, on datapoint 2000, the information overload is not yet high that the actual tweet behaviour \(X_0\) is entirely suppressed even though the individual is receiving much information. This behaviour makes sense as they are in line with what one would expect from the new individual behaviour.

\(^2\) [https://textblob.readthedocs.io/en/dev/](https://textblob.readthedocs.io/en/dev/)
For the following experiments, on top of the basic model, some parameters are adjusted using simulated annealing as described above. The training points for the algorithm were based on the actual behaviour of a user on Twitter. The observed behaviour is quite different from the two earlier described models. Fig. 5 shows the best model with an RMSE of 0.06667. With this set of values, it can be observed that the overall information overload is quite high, but also the offline interactions are quite high.

The model both has a high persistence factor, and a low speed factor as a result of a strong persistence overall. It can be observed that the stimuli of new tweets have a more considerable influence in the dropout or fed-up factor but tend to decrease as much as it increases due to these stim-

Table 1
Important data features extracted.

| Feature   | Description                                      |
|-----------|--------------------------------------------------|
| Original Author | Name of the original author                      |
| Is Source | User is a twitter verified account               |
| User Mentions | Mentions in the tweet                            |
| Hashtags  | Hashtags in the tweet                            |
| Sentiment | Sentiment of the twitter text [0–1].             |
| Engagement | Number of tweets talking about topic over 50 last tweets |
| Relationship | Represents if the user follows the source [0–1]  |

Fig. 2. Time interpolation for each relevant tweet from a source.

Fig. 3. Basic model while experiencing overload. One can see that as soon as the information overload exceeds the fed-up stimuli, the actual Twitter behaviour gets suppressed to almost zero. As an effect, someone has such an information overload that she or he thus stops tweeting about the topic.
Basic model with reduced information overload

Fig. 4. Basic Model with reduced information overload.

Input and Target based on actual Twitter engagement

Input data and resulting model based on actual Twitter engagement

Fig. 5. Input data and resulting model based on actual Twitter engagement.
Thereby, there is barely any Twitter interaction during the model as everything is condensed just after receiving the stimuli. Over time, the offline interactions decrease faster than the information overload, and thus it results in a very slightly decreasing Twitter engagement tendency when new stimuli come. The best way to fit the behaviour of the individual is by assuming that there is a more extended history than the two week time span of the dataset is used for this paper. Accordingly, one observes that the initial values of the information overload and offline interactions already start both high.

Finally, some data points were created to model different user engagement, and based on those data points, the same parameters were tuned using simulated annealing.

In this case, the simulated annealing produced a set of new parameter values with an RMSE of 0.4116. The behaviour of the model with the new set of parameters can be observed in Fig. 6. This experiment resulted in a behaviour that was closer to the behaviour observed in the basic models explained above. The described behaviour best suits an individual that already has been engaged in the topic and has experienced information overload. As an effect, in the beginning, the user is barely tweeting about the subject, ignoring new stimuli. However, over time as the information overload has decreased, new stimuli are registered and result in some quick Twitter engagement after which the information overflow increases again, reducing future interactions for a given period. Once the fed-up factor has decreased again, the individual shows some, but reduced, new Twitter behaviour on new stimuli again.

5. Verification of the model by mathematical analysis

This section details an analysis of the stationary points for the maxima and minima in the final stage for the important states plotted in the simulations (X6, X7, X8, X9). In particular, stationary points occur for these states in the two simulations addressed below. The first one refers to Fig. 3 and is shown in Table 2.
The results are consistent with the simulations performed above confirming the mathematical accuracy of the model for the original model as all the values of the absolute deviation are lower than $10^{-2}$. One can observe that at the moment that the information overload state $X_5$ has a high value, the other states are stable around their lowest possible value given the alogistic combination function. The information overload state becomes stable after a long period of no new stimuli from new tweets. However as seen in Fig. 3, the state becomes instable again as soon as there are new stimuli, which are not suppressed due to the low information overload.

Finally, Table 3 shows the mathematical analysis for the last simulation, illustrated in Fig. 6. For this case the simulation shows a behaviour similar to the behaviour described for the first mathematical analysis where all the states have a deviation smaller than $10^{-2}$, and thus confirming the mathematical accuracy of the implemented model.

6. Discussion

The current paper proposes an adaptive temporal-causal network to study the process of sharing information and simulate the impact of information overload in such processes. The network model design is based on the ideas of other authors (Boyd et al., 2010; Recuero et al., 2011; Suh et al., 2010) that identify certain features as necessary for the information sharing process. In addition to them, this model adds two other essential features that are usually ignored, the offline behaviour and information overload. These two extra features help to describe better how the actual information persists over time and how individuals are willing to share, depending on how much information is already flowing freely.

The four reported simulations describe different behaviour patterns depending on the used parameters for the two newly introduced factors. The three settings that modelled fictitious individuals successfully describe the prior belief of individual behaviour. Furthermore, it proved to be flexible to model individuals with different tolerances for information overflow. The model that was fitted on the actual behaviour of a twitter user turned out a bit different than the other models. Although the results could be a proper simulation of the actual behaviour, it could as well be that those are related to the used methods. As can be seen in the input data, the observed datapoints have quite some covariance with the input data, which could reduce the generalisation performance of the algorithm. Another possible explanation could be that the heuristic function has been overfitting on this data, as it created a rather low RMSE.

The proposed model can be extended in several ways. Firstly, the dataset needs to be larger, both in queries affecting a topic and in the number of rows (tweets) the dataset contains. Also, the time span of two weeks of the database turned out to be rather short. A larger dataset capturing more information over time is more useful and will allow obtaining more information on the long term processes. A larger and richer dataset will remove some of the theoretical assumptions made about individuals making the full model more robust and valid in the real world. Another interesting addition for future research is to explore the use of a regularisation factor to increase the generalisation performance of the model.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Bonnett, W. (2019). Scottish independence: The vote that could destroy the United Kingdom. Retrieved December 24, 2019, from CNN website: https://edition.cnn.com/interactive/2019/12/world/uk-election-scottish-independence/.

Boyd, D., Golder, S., & Lotan, G. (2010). Tweet, tweet, retweet: Conversational aspects of retweeting on Twitter. In Proceedings of the 43th Annual Hawaii International Conference on System Sciences (pp. 1–10). https://doi.org/10.1109/HICSS.2010.412.

Chewning, E. G., & Harrell, A. M. (1990). The effect of information load on decision makers’ cue utilization levels and decision quality in a financial distress decision task. Accounting, Organizations and Society, 15(6), 527–542. https://doi.org/10.1016/0361-3682(90)90033-Q.

CNN (2019). Hong Kong protests: Updates and latest on city’s political unrest - CNN Retrieved December 24, 2019, from https://edition.cnn.com/specials/asia/hong-kong-protests-intl-hnk.
Ellison, N.B., & Boyd, D.M. (2013). Sociality through social network sites. In: W.H. Dutton (Ed.), The Oxford Handbook of Internet Studies (pp. 151-172). https://doi.org/10.1093/oxfordhb/9780199589074.001.0001.

Eppler, M. J., & Mengis, J. (2004). The concept of information overload: A review of literature. The Information Society, 20(5), 325–344. https://doi.org/10.1080/01972240490507974.

Goel, S., Watts, D. J., & Goldstein, D. G. (2012). In The structure of online diffusion networks (pp. 623–638). New York: ACM Press. https://doi.org/10.1145/2229012.2229058.

Gomez-Rodriguez, M., Gummadi, K. P., & Schölkopf, B. (2014). In Quantifying information overload in social media and its impact on social contagions (pp. 170–179). AAAI Press.

Gomez-Rodriguez, M., Leskovec, J., & Krause, A. (2012). Inferring networks of diffusion and influence. ACM Transactions on Knowledge Discovery from Data, 5(4). https://doi.org/10.1145/2086737.2086741.e21.

Hill, S. J., Lo, J., Vavreck, L., & Zaller, J. (2013). How quickly we forget: the duration of persuasion effects from mass communication. Political Communication, 30(4), 521–547. https://doi.org/10.1080/10584609.2013.828143.

Independencia de Cataluña | EL MUNDO. (2019). Retrieved December 24, 2019, from El Mundo website: https://www.elmundo.es/t/in/independencia-catalunya.html.

Kim, J., & Yoo, J. (2012). In Role of sentiment in message propagation: Reply vs. retweet behavior in political communication (pp. 131–136). IEEE Explore. https://doi.org/10.1109/SocialInformatics.2012.33.

Lee, M., Kim, H., & Kim, O. (2015). Why do people retweet a tweet?: Altruistic, egoistic, and reciprocity motivations for retweeting. Psychologia, 58(4), 189–201. https://doi.org/10.2117/psy soc.2015.189.

Lewis, A., & Perez Maestro, L. (2019). Catalan protests: President Quim Torra calls for negotiations with Spanish government after fifth night of violent unrest - CNN. Retrieved December 24, 2019, from CNN website: https://edition.cnn.com/2019/10/19/europe/barcelona-catalonia-protests-torra-negotiations-intl/index.html.

Liu, Y., & Ying, X. (2012). A review of social network sites: definition, experience and applications. In Proc. of the Conference on Web Based Business Management (pp. 749–752).

Mitchell, A., Rosenstiel, T., & Christian, L. (2012). What Facebook and Twitter mean for news. In Pew Research Center. Retrieved from http://stateofthemedia.org/2012/mobile-devices-and-news-consumption-some-good-signs-for-journalism/what-facebook-and-twitter-mean-for-news/?src=proc-section.

Recuero, R., Araujo, R., & Zago, G. (2011). How Does Social Capital Affect Retweets? Proc. of the Fifth International AAAI Conference on Weblogs and Social Media, ICWSM'11 (pp. 305-312). AAAI Press.

Reuters. (2019). Timeline: Key dates in Catalonia’s independence bid - Reuters. Retrieved January 5, 2020, from https://www.reuters.com/article/us-spain-politics-catalonia-timeline/timeline-key-dates-in-catalonias-independence-bid-idUSKBN1WS0S9.

Schoenmaker, R., Treur, J., & Vetter, B. (2018). A temporal-causal network model for the effect of emotional charge on information sharing. Biologically Inspired Cognitive Architectures, 26, 136–144. https://doi.org/10.1016/j.bica.2018.10.003.

Suh, B., Hong, L., Pirolli, P., & Chi, E. H. (2010). Want to be retweeted? Large scale analytics on factors impacting retweet in twitter network. In Proceedings of SocialCom 2010: 2nd IEEE International Conference on Social Computing (pp. 177–184). https://doi.org/10.1109/SocialCom.2010.33.

Tighe, P. J., Goldsmith, R. C., Gravenstein, M., Bernard, H. R., & Fillingim, R. B. (2015). The painful tweet: Text, sentiment, and community structure analyses of tweets pertaining to pain. Journal of Medical Internet Research, 17(4). https://doi.org/10.2196/jmir.3769.e84.

Treur, J. (2016). Network-Oriented Modeling: Addressing Complexity of Cognitive, Affective and Social Interactions. Cham, Switzerland: Springer Nature Publishers.

Treur, J. (2020). Network-Oriented Modeling for Adaptive Networks: Designing Higher-Order Adaptive Biological, Mental and Social Network Models. Cham, Switzerland: Springer Nature Publishers.