

Information Management System

Technical Field

The present invention relates, in general terms, to an information management system, and also relates to methods of managing information.

Background

Recently, media and reports have been increasingly digitalized. Electronic devices such as IPAD, IPHONE, Tables and PCs have been introduced so that users can read news and information on line or in a digital form. Moreover, ads and clips have been increasingly displayed on line.

All the existing applications provide common features such as a UX/convenience, inter-operability, design, data security, and price. However, all these applications do not offer a way of debiasing the information before it is provided to the client. All the applications are passive data repositories. They allow individual usage or team usage but just as a passive platform. At the same time, many sources of news (Newspapers, news Analytics) provide stream of data delivered in a sequential way. The problem in both cases is that the user of the data is subject to several behavioural biases depending on how the data/information is presented.

The way the pieces of news are displayed has major implications for the way users react to them. Indeed, it has been shown that the sequencing of news of different tone, the length of the sequence, the degree by which different tones are spread over time, the positioning within the sequence has different effects on individual perception and mood. Also, it affects people's ability to process the information contained in the sequence, the willingness to take action, the degree of risk aversion. This has major implications in terms of mood of the user as well as his ability to function effectively.

It would be desirable to overcome all or at least one of the above-described problems.

Summary

Disclosed herein is a computer-implemented method of presenting a sequence of information items to a user, the method comprising:

- determining a behavioural profile and/or a behavioural objective;

- for each of the sequence of information items, performing a sentiment analysis to determine a sentiment score for the respective information item, the sentiment scores for all information items forming a set of sentiment scores for the sequence of information items;

- determining whether the set of sentiment scores is compatible with the behavioural profile and/or the behavioural objective;

- responsive to a determination that the set of sentiment scores is sub-optimal in view of the behavioural profile and/or the behavioural objective, resequencing the sequence of information items to optimize the sequence of information in view of the behavioural profile and/or the behavioural objective; and

- causing the sequence of information items to be presented to the user via a display of a user device.

In some embodiments, the behavioural objective is a mood of the user, a risk-taking capacity of the user, a change in attention bias of the user, or a difference in attitude of the user to one or more other users.

In some embodiments, said resequencing comprises modifying one or more of: the order of the information items; temporal spacing between successive information items; and sub-groupings of the information items within the sequence.

In some embodiments, said resequencing comprises adjusting temporal spacing between successive information items within a user-allocated viewing time period.

In some embodiments, the set of sentiment scores is determined at least partly based on metadata associated with the information items.

In some embodiments, resequencing comprises resequencing the information items so that the set of sentiment scores is compatible with the behavioural profile and/or the behavioural objective.

Disclosed herein is also a system of presenting a sequence of information items to a user, the system comprising a plurality of processors configured for:

- determining a behavioural profile and/or a behavioural objective;

- for each of the sequence of information items, performing a sentiment analysis to determine a sentiment score for the respective information item, the sentiment scores for all information items forming a set of sentiment scores for the sequence of information items;

- determining whether the set of sentiment scores is compatible with the behavioural profile and/or the behavioural objective;

- responsive to a determination that the set of sentiment scores is sub-optimal in view of the behavioural profile and/or the behavioural objective, resequencing the sequence of information items to optimize the sequence of information in view of the behavioural profile and/or the behavioural objective; and

- causing the sequence of information items to be presented to the user via a display of a user device.

Brief description of the drawings

Embodiments of the present invention will now be described, by way of non-limiting example, with reference to the drawings in which:

Figure 1 is a flow diagram showing a method used for presenting a sequence of information items to a user; and

Figure 2 is a schematic diagram showing components of an exemplary computer system for performing the methods described herein.

Detailed description

The present invention relates to a method of sequencing the ordering of a set of news and a reading and selecting apparatus for performing the method. The focus of the present disclosure is on the *order of news*. The sequence of news can be ordered to represent an improving or worsening trajectory, or may lack any discernible trajectory, for example, a flat line or fluctuating performance.

The present disclosure also relates to a method for reducing the distorting effect produced by the sequencing of information. It will be appreciated that the way information is provided affects people's behavior. For example, the sequencing between news of different tone (i.e., "good" and "bad" news) affects the attitude towards risk, distorts the reference point of the user, anchors its behaviour, and hijacks the user's attention. The proposed method eliminates such distortions or use them to the user's advantage. It interfaces with the information that is provided and reshuffle it putting in a sequencing that is useful to the user.

Disclosed herein is a method of sequencing the ordering of a set of information and a system for performing the method. The method includes going over the information provided by a series of outlets, identifying the sentiment/mood of such information and sequencing the display of the information according to an algorithm meant to maximize the goal of the user. In general, the present invention is about the usage of techniques of textual analysis to read the information, extract the sentiment and sequence the provision of the news.

Said method operates in several parts. In the first part, an objective is determined. The objective may be either minimizing the emotional negative impact of the displayed information or improving the decision process that exploits such information. The definition of the information is very broad. It may cover both specific information in specific contexts, e.g., information provided before a meeting as well as information required by a single decision-maker. The information may also cover contents provided by media outlets, e.g., media news, digital newspapers, videos, digital broadcasting.

In the second part, the method reads the information and do a textual analysis of it to identify the mood of the information, that is, whether the information is “good” or “bad”. In one embodiment, the proposed method involves deciphering the sentiment contained in the information to determine whether the information is “good” or “bad”. It will be appreciated that the way the pieces of information are displayed has major implications for the way users react to them. Indeed, it has been shown that the sequencing of information of different tone, the length of the sequence, the degree by which different tones are spread over time, the positioning within the sequence, has different effects on individual perception and mood. Also, the way the information is displayed affects people’s ability to process the information contained in the sequence, the willingness to take action, the degree of risk aversion. This has major implications in terms of mood of the user as well as his ability to function effectively.

In the final part, the application will reposition the different pieces of the information according to a behavioural sequence that maximizes the goal of the first step, making them available to the user. The present disclosure aims to sequence the provision of the information as a function of such sentiment with the goal of achieving specific objectives. The sequence technique is proprietary and is designed to optimize a goal function of the user (e.g., happiness, effectiveness). In some embodiments, an optional final step can be added that provides an indicator of the general mood of the day.

Figure 1 illustrates an example computer-implemented method 100 of presenting a sequence of information items to a user, the method 100 comprising:

Step 102: determining a behavioural profile and/or a behavioural objective;

Step 104: for each of the sequence of information items, performing a sentiment analysis to determine a sentiment score for the respective information item, the sentiment scores for all information items forming a set of sentiment scores for the sequence of information items;

Step 106: determining whether the set of sentiment scores is compatible with the behavioural profile and/or the behavioural objective;

Step 108: responsive to a determination that the set of sentiment scores is sub-optimal in view of the behavioural profile and/or the behavioural

objective, resequencing the sequence of information items to optimize the sequence of information in view of the behavioural profile and/or the behavioural objective; and

Step 110: causing the sequence of information items to be presented to the user via a display of a user device.

The method 100 may comprise using a feeder to download or feed in a system for processing the method 100 different types of information coming from different sources. One example of information is made of news. The system for processing the method 100 may interconnect with the source of news (e.g., FT, WSJ, NYT) downloading the articles, which is the first set of news. The second set of news is made of briefs and reports provided before a meeting (e.g. board meeting). The third set of news are videos provided in clips or streaming.

At Step 102, an interface may be used to feed in the behavioural profile and/or behavioural objectives. In some embodiments, said behavioural profile is recorded as the system for processing the method 100 is first used. It may contain information on major characteristics of the user, such as age, wealth, attitude towards risk, etc. The behavioural objectives may be recorded as the system for processing the method 100 is activated and include a selection out of a menu of the objective to be achieved. In some examples, it is the user who determines the behavioural objectives of the system for processing the method 100.

In the present disclosure, there are various kinds of behavioural objectives. It will be appreciated that those behavioural objectives are not quantified in the present disclosure. The first kind is related to happiness. In other words, the sequencing of information is created in order to reduce unhappiness. When the user finishes using the system for processing the method 100, he is expected to have the information he needs with the smallest possible discomfort.

The second kind is related to risk taking. Risk taking is an object terms described in the existing literature. In the present invention, the sequencing of news is created in order to reduce the distortion to the ability of the user to take decisions that involve risk. When the user finishes using the system for

processing the method 100, he is expected to have the information he needs in a form that helps him to optimize his risk taking.

The third kind is related to anchoring/availability/attention bias. It will be appreciated that the present disclosure relies on the existence literature which describes the existence of biases. In the present invention, the sequencing of information is created in order to reduce distortion in the decision process due to changing level of attention created by the wrong sequencing of information. When the user finishes using the system for processing the method 100, he will have the information he needs with the smallest possible attention bias.

The fourth kind is related to brainstorming and herding. The system for processing the method 100 may be used in a group/team context (e.g., board meeting) with the purpose of creating heterogeneity in the group decision making. The sequencing of information may be created in order to increase heterogeneity in the response to information. In one example, at the beginning of the meeting, the group members will have got the information from the application. The system for processing the method 100 will provide each member of the group a different reference/anchoring point and with different mental frame and attitude towards risk. The goal in this case is to help framing the problem in a new way exploiting heterogeneity within the group.

At Step 104, a browsing apparatus is used to quickly grasp the metadata contained in the information. Said metadata contains the tone of the information, its emotional and behavioural context, the interest and perception by the user, the degree of attention arousal, as well as the overall hedonic profile.

The browsing apparatus will include a mechanism of textual analysis meant to extract and quantify the metadata of the information. The operation of interpreting the metadata includes the operations of extracting the words, symbols and images from news and videos for each of the selected information, and analysing them using textual analysis to extract the tone of the selected information. Such mechanism will further include the operation of quantifying

specific feature of the metadata related to the emotional and behavioural profile of the user.

For example, in the case of tone, the information is read by the application and processed. A textual analysis algorithm will read the news and identify the tone of it. It will then provide an indicator of it (e.g., 0 = very gloomy, 50 = average, 100 = very optimistic and good). Different types of algorithms can be envisaged depending on whether the type of information (a) mostly words, b) mostly numbers, c) mostly visual)).

Once the degree of sentiment of the information has been identified, all the information in the set (e.g., block) will be ranked from the ones with the highest positive sentiment to the ones with lowest negative sentiment.

Some examples of textual analysis methods are discussed below.

The first kind of textual analysis method is neural network model for sentiment analysis. The first example is convolutional neural network (CNN). Let $X = (x_1, x_2, \dots, x_i, \dots, x_n)$ represent the sequence of words of a business news article, where x_i is a k -dimensional vector encoded to represent i -th word in the article. We train a model $y = F(X)$, where y is the sentiment of the news article, and F is a mapping from X to y . The present invention imposes a convolutional neural network (CNN) architecture on F to incorporate the fact that the sequence of words is at least as important as occurrence of an individual word. In CNN, in order to capture the information contained in a sequence of words, the present disclosure transforms a sequence of words as new features instead of feeding X directly into a fully connected neural network.

Specifically, let $X_{i,i+h} = (x_i, \dots, x_{i+h})$ be sub-sequence of words from X . The present disclosure defines $c_i = f(w \times \hat{X}_{i,i+h} + b)$, where w and b are the parameters (to be estimated) of a linear transformation, $\hat{X}_{i,i+h}$ is the flattened vector by concatenating $X_{i,i+h}$, and f is a hyperbolic tangent function. The transformation is applied to each possible sub-sequence in X for each $i \in \{1, \dots, n - h + 1\}$. Then we further get an input feature by pooling $\hat{c}_h = \max(c_h)$,

where $c_h = [c_1, \dots, c_{n-h+1}]$. In addition, we set h to different values to get $C = [\hat{c}_{h_1}, \dots, \hat{c}_{h_H}]$. We then feed the vector C as the input to a fully connected neural network. y is defined as the normalized stock returns during the time window $[-1, 1]$, i.e., the previous, current and next day around the news article, where we normalize the stock returns by subtracting the mean and divided by the standard deviation of news event returns of the same trading day.

Computationally, we start the CNN model with an embedding layer on the top. We use GloVe for the embedding. Then three convolutional layers are stacked with the filter widths of 3, 4, and 5. The convolutional layers are followed by a 1-Dimensional maximum pooling. Then the present disclosure applies dropout=0.5 before finishing the architecture with a fully connected linear layer with the output dimension to be 1. When we train the model, we use the mean squared error as the loss function, and adam as the optimizer which derives from adaptive moment estimation. The model $y = F(X)$ for year yr is trained with all the news articles occurred from year $yr - 5$ to year $yr - 1$. We randomly split 70% of the text data as the training dataset and keep the remaining 30% for the validation. We train the model for 20 epochs, and keep the best performing model (measured by the validation loss).

The second example is recurrent neural network (RNN) with gated recurrent units (GRUs). RNN uses a recurrent hidden state to record the information contained in the previous items in a sequence of words (e.g., news articles). Specifically, given a sequence $x = (x_1, \dots, x_t, \dots, x_T)$, RNN update its hidden state $h_t = \phi(h_{t-1}, x_t)$ for $t > 0$. Intuitively, h_t keeps a record of previous information (i.e. h_{t-1}) and the current information (i.e. x_t), and carrying forward, h_t is used to update h_{t+1} . Traditionally, ϕ is implemented as $h_t = g(Wx_t + Uh_{t-1})$, where g is a logistic sigmoid function or a hyperbolic tangent function that takes in an affine transformation of x_t and h_{t-1} as the input.

Unfortunately, it is difficult to keep tracking h_t with a vanilla RNN architecture for a stable estimation because of a gradient vanishing problem. The problem emerged when h_t keep track a long chain of previous words, and this problem can be alleviated by introducing a smarter gating unit to record some information and forget the rest of the information. GRU can outperform LSTM

both in terms of convergence and generalization. Therefore, in the present disclosure, we will adopt GRUs in our RNN sentiment analysis model.

Specifically, j^{th} element of h_t is computed as $h_t^j = (1 - z_t^j)h_{t-1}^j + z_t^j\tilde{h}_t^j$, where $z_t^j(x_t, h_{t-1})$ decide the extent of h_t^j will its content, and $\tilde{h}_t^j(x_t, h_{t-1})$ is the update of the contents. $z_t^j(x_t, h_{t-1})$ is computed by $\sigma(W_z x_t + U_z h_{t-1})^j$, where σ is a logistic sigmoid function and W_z , and U_z , which will be estimated during the model training, are matrix to have an affine transformation of x_t and h_{t-1} . The updated contents $\tilde{h}_t^j(x_t, h_{t-1})$ is estimated as $\tilde{h}_t^j = \tanh(W x_t + U(r_t \odot h_{t-1}))$, where $\odot$ is an element-wise multiplication operator, and $r_t^j = \sigma(W_r x_t + U_r h_{t-1})^j$ is called reset gate which determine how much of the previous state to forget.

Operationally, we start a GRU-RNN model with an embedding layer of GloVe. Then we add a two-layer of bi-directional GRUs, where we realize the bi-directional units by feeding both the original sequence and the reversed sequence to the GRUs layer. We apply a dropout=0.25 before finishing the network with a densed fully connected linear layer with output dimension to be 1. Figure # illustrates the architecture. When we train the model, we use the mean squared error as the loss function, and adam as the optimizer. We train the model for 20 epochs, and keep the best performing model.

The third example of the neural network model for sentiment analysis is neutral network with a bag of N-gram tricks (fastText). The bag of n-grams to capture some partial information embedded into a sequence of words. N-gram is a contiguous sequence of n items from a given sample of text or speech. For example, for a sentence (EPS, beats, market, expectation), the collection of all the 2-grams is (EPS beats, beats market, market expectation). We follow the above approach to build a neutral network model for sentiment analysis. Let $W = (w_1, \dots, w_m)$ be a sequence of words (i.e. a news article). The features feed into a neutral network is a sequence of n-gram instead each individual word. Specifically, let $X_n = (x_1, \dots, x_i, \dots, x_{m-n+1})$, where x_i is k-dimensional vector encoded to represent an n-gram $w_i w_{i+1} \dots w_{i+n-1}$. We build input features $X = (X_1, \dots, X_N)$ by taking the collection up to N-grams. Then we train a model $y =$

$F(X)$, where y is the sentiment of the news article, and F a fully connected neural network with a flattened vector of X as the input.

The architecture of neural network with N-gram tricks in the present disclosure is rather straightforward. First, the original text $W = (w_1, \dots, w_m)$ is expanded into an augmented sequence with 2-gram and 3-gram appended to the end of the original sequence. Then the present invention feeds w_{new} into a simple neural network with an embedding layer followed by a fully connected linear layer. When training the model, the present invention uses the mean squared error as the loss function, and adam as the optimizer. We train the model for 20 epochs, and keep the best performing model.

The fourth example of the neural network model for sentiment analysis is BERT which is the first language model enable transfer learning. BERT frameworks consist of two steps: pre-training and fine-tuning. During pre-training, the model is trained on huge number of unlabelled data with huge number of parameters. Then the model after the pre-training can be used for various downstream natural language task through fine tuning. BERT is revolutionary because it is a unified architecture cross different task. Moreover, BERT shoulders the computational expense steps in the pre-training, and then the downstream user of the model will have a light weight fine tuning task with an affordable computational resource, making it possible for someone without cutting-edge computer to develop cutting-edge AI application. In the present disclosure, we use the pre-trained BERT model provided by Google to build a sentiment analysis model with a fine-tuning step.

BERT's power partially arise from its input representation. BERT encodes a word in text with three dimensions. First, BERT has a token embedding to represent a word itself. Second, BERT has a segment embedding to represent the sentence that a word belongs to. Third, BERT has a position embedding to represent the relative position of the word in each sentence. With the three-dimensional vector representation of a given word, BERT does not only model what a given word means, but also model the context of the word (i.e., which sentence this word shows up, and the relative position of the word in the sentence). It is very

straightforward to apply BERT to a downstream task modelling and inexpensive to compute.

Operationally, the present invention incorporates the pre-trained BERT model into the sentiment analysis model. We start our BERT model with a BERT layer on the top. Similarly to the CNN mode, three convolutional layers are stacked with the filter widths of 3, 4, and 5. The convolutional layers are followed by a 1-Dimensional maximum pooling. Then dropout=0.5 is applied before finishing the architecture with a fully connected linear layer with the output dimension to be 1. Due to the powerful embedding representation of BERT, we reduce the number of epochs to 5.

The second kind of textual analysis method is dictionary-based sentiment analysis model. In the prior art, Harvard-IV psychological dictionary is used to estimate the sentiment. The pessimism category predicts the stock market returns. Harvard-IV does not specific to the context of finance related language. A dictionary was used in the prior art to fit the financial context. However, both dictionaries did not consider jointly the words in the text. For example, the GI (General Inquirer) is a pure word count program, so it does not categorize combinations of words that often possess different meanings from the constituent words. *The* present invention conducts a sentiment analysis with both Harvard-IV and Loughran and McDonald (ML) dictionary. Specifically, we use the positive and negative categories in Harvard-IV and ML dictionary. The present disclosure counts the number of positive and negative words in each article and define the sentiment of the article as $\frac{\#Positive\ Words - \#Negative\ Words}{\#Positive\ Words + \#Negative\ Words}$.

The third kind of textual analysis method is named sentiment extraction via screening and topic modelling (SESTM). In particular, a statistical-based sentiment analysis framework is used without relying on pre-existing dictionaries. The occurrence of the sentiment-charged words in news article i is modelled as a mixture multinomial distribution $Multinomial(s_i, p_i O_+ + (1 - p_i) O_-)$, where s_i is the total counts of sentiment-charged words in article i and determines the scale of the multinomial. O_+ is the expected frequencies in a

maximally positive sentiment articles, and O_- is the word frequencies in maximally negative articles. Let $F = \frac{1}{2}(O_+ + O_-)$ and $T = \frac{1}{2}(O_+ - O_-)$.

Intuitively, if a word has a larger value in F , then the word appears more frequently overall. If a word has a larger value in T , its sentiment is more positive. p_i is the probability that the article i has a positive sentiment. Numerically, the sentiment of an article is p_i such that p_i maximize the likelihood of occurrence of the sentiment charged words in the article i . Specifically, the estimated $\hat{p} = \arg \max_{p \in [0,1]} \left\{ \frac{1}{S} \sum_{j=1}^S d_j \log(p \hat{O}_{+,j} + (1-p) \hat{O}_{-,j}) + \lambda \log(p(1-p)) \right\}$, where d_j , $\hat{O}_{+,j}$ and $\hat{O}_{-,j}$ can be estimated from data, and $\lambda \log(p(1-p))$ is a penalty term for the accuracy guarantee.

SESTM is implemented in three steps. The first step is the screening for sentiment-charged words. Words are ranked by $f_j = \frac{\text{\#articles with word } j \text{ AND stock return is positive}}{\text{\#articles with word } j}$, and then the top and bottom n words are used as the sentiment charged words and drop the rest of the words.

A grid search is done for $n = [25, 50, 100]$ and n that perform the best in the test sample during the training is chosen. The second step is the topic modelling where d_j , $\hat{O}_{+,j}$ and $\hat{O}_{-,j}$ are estimated on the training dataset. In step 3, the sentiment of a new article is estimated by $\hat{p} = \arg \max_{p \in [0,1]} \left\{ \frac{1}{S} \sum_{j=1}^S d_j \log(\hat{p}_i \hat{O}_{+,j} + (1 - \hat{p}_i) \hat{O}_{-,j}) + \lambda \log(p(1-p)) \right\}$.

The fourth kind of textual analysis method is named explainable AI (XAI). An integrated gradient along x_i , the i -th dimension of the input of x , as $IntegratedGrads_i(x) = (x_i - x'_i) \times \int_{\alpha=0}^1 \frac{\partial F(x' + \alpha \times (x - x'))}{\partial x_i} d\alpha$. Computationally, they approximate the integrated gradient with a finite number of Riemman approximation of the integral: $IntegratedGrads_i^{approx}(x) = (x_i - x'_i) \times \sum_{k=1}^m \frac{\partial F(x' + \frac{k}{m} \times (x - x'))}{\partial x_i} \times \frac{1}{m}$, where numerically $\frac{\partial F(\cdot)}{\partial x_i}$ can be inexpensively obtained with the mainstream deep learning libraries. In some example, we can obtain the

gradient with torch.autograd.grad efficiently in Pytorch library. Tensorflow library also provides the utility tf.GradientTape to quickly compute the gradient.

The output gives an attribution score ($IntegratedGrads_i^{approx}$) to each word in the business news. The method makes a neural network as transparent as the dictionary-based method and statistical methods. Dictionary based methods is transparent because we know which words are important in driving the sentiment score. We can obtain the important words by looking at words with the highest/lowest IG score. Statistical-based SSESTM is transparent because we know the list of the sentiment-charged words and the positive and negative topic coefficient of each keyword. The IG score is comparable to the topic coefficients in SSESTM model. For a given article, the words among those with the highest/lowest IG scores would be the words heavily charged with sentiment towards to overall sentiment of the articles.

With Integrated Gradient (IG) score, the present invention is able to identify the words with the highest power to determine the sentiment of an article. The present disclosure defines the sentiment score of a sentence as the maximum IG of all words in the sentence. With the sentence level score, we can visualize what sentence is driving the overall sentiment of the news article.

IG score of the word will be varying depending on the context of the word. We can identify what words are driving the IG of a given word with the following procedure. In a sequence of words $(w_1, w_2, \dots, w_i, \dots, w_m)$, we identify the words that affect IG of word w_i with the following procedure. First, we collect the words $(w_{i-10}, \dots, w_{i-1}, w_{i+1}, \dots, w_{i+10})$ for each article that contains word w_i , which are the adjacent words that might affect the sentiment charged by w_i . Second, we run a screening process to identify the words with higher probability to show up around w_i . Specifically, we rank adjacent words w_j by $\log\left(\frac{\#w_j \text{ adjacent to } w_i}{\#words \text{ adjacent to } w_i}\right) - \log\left(\frac{\#w_j \text{ in all texts}}{\#words \text{ in all texts}}\right)$, and keep the top n words as the candidates. Then we calculate the impact of w_j on w_i with a simple regression $IG_{i,a} = \beta_{i,j} \times I_{w_j \in (w_{i-10}, \dots, w_{i-1}, w_{i+1}, \dots, w_{i+10}) \text{ in } a} + \beta_{i,j,0}$, where $IG_{i,a}$ is the integrated gradient score of word w_i in article a , $I = 1$ if w_j is adjacent to w_i in the article a . We define w_j as a catalyst word to w_i if T-statistics of $\beta_{i,j}$ is above 2 or below -2.

At Step 108, one or more algorithms are selected to sequence the information according to the objectives of the user using the metadata. Step 108 refers to the optimizing the sequence of the information. It will be appreciated that the term "optimizing" and similar refer to improving the sequence of information items in view of the behavioural profile and/or behavioural objective. This may result in an optimal (i.e. best) sequence of the information items, or will at least result in an improved sequence that more closely aligns with the behavioural profile and/or behavioural objection. As will be discussed in detail, resequencing may mean reordering the items - e.g. most negative sentiment to most positive sentiment for a profile seeking a happy feeling - or may involve other processes such as substituting one news item for a news item on the same topic but with different sentiment score.

The algorithm applied in step 108 exploits behavioural techniques that choose the sequencing in order to achieve specific objectives selected by the users in the interface (e.g., increase happiness, increase effectiveness). The general principle underlying the resequencing of the pieces of information is defined in accordance with the user's objective. In some embodiments, the goal is to exploit the metadata to create a "behavioural sequencing" geared at maximizing the objectives of the user given his behavioural profile.

In one example, when the objective is to increase happiness, the sequencing of the pieces of information will be from the ones with more negative sentiment to the ones with more positive sentiment. That is, the information is sequenced as a function of the index of sentiment. The one with the lowest sentiment first, then the one with the next to lowest sentiment up to the piece of information with the highest degree of sentiment. In another example, the objective is to increase risk taking, and in such case the sequencing of the pieces of information will be from the ones with positive sentiment to the ones with more negative sentiment. That is, the information is sequenced as a decreasing function of the index of sentiment. The one with the highest sentiment first, then the one with the next to highest sentiment up to the piece of information with the lowest degree of sentiment.

When the objective is to reduce anchoring/availability/attention bias, the sequencing of the pieces of information is created in order to reduce distortion in the decision process due to changing level of attention created by the sequencing of information. The distortion has been shown to be related to the fact that a high-impact piece of information reduces attention for the other synchronous provisions of information.

For example, agents and consumers tend to judge a product's quality following different sequences. Agents normally evaluate multiple sequential cues of varying strengths in order to draw inferences about a product's quality. When they are not distracted, they follow a strong-weak cue sequence relative to a weak-strong sequence (i.e., a primacy effect). However, if they are distracted, the primacy effect reverses to a recency effect. On the contrary, consumers judge a product's quality more favourably following a weak-strong cue sequence (i.e., a recency effect). This implies that high-impact information may induce recency and induce forgetting valuable information. This can be reduced in two ways. The first way is positioning high impact information first so that their impact fades away. At the same time, maintaining a stable level of sentiment helps to reduce the attention deficit.

The present invention therefore proceed as follows. At step one, we first identify all the pieces of information with extreme degree of sentiment (i.e., very high positive sentiment and very high negative sentiment) and then we position them at the beginning of the sequence. At step two, we alternate positive and negative sentiment information so as to minimize the overall "impression" and keep an average level of sentiment. This is done by calculating the average level of sentiment across the different pieces of information and minimizing the variation of it in the sequencing. To minimize the variation, we minimize the absolute difference between the cumulated level of sentiment and its average level. It will be appreciated that step one is preliminary and step two is conditional on step one.

In the present disclosure, when the objective is to increase brainstorming/herding, the sequencing of news will be created in a different way for each participant to maximize the absolute level of cumulative sentiment

variation among the participants. In particular, each participant will receive a different sequencing. Each sequencing will have a specific degree of cumulative sentiment constructed as follows. Each piece of information receives a decreasing weight as a function of how late in the sequencing process the piece of information is. For example, in the case of 10 pieces of information, the first will receive a weight of 1, the second of 2, the third of 3 and so on. The cumulative weighed average of the different pieces of information represents the degree of information intensity of the sequence. Then, the algorithm will select the sequencing for each participant such that the absolute variation in the degree of information intensity across participants is the highest.

Referring to the objective set of the user, an algorithm will sequence the information according to key quantifiable inputs from the metadata. The operation of sequencing the information is performed by referring to an algorithm that chooses the order of displaying of the information, the space between them, the grouping of information/news/videos in broader or finer sections and sub-sequences, the trend, the hedonic profile of the sequence and sub-sequences, the evenness of their spreading within the time frame the user allocates to the perusal/view.

The present invention also relates to a system of presenting a sequence of information items to a user, the system comprising a plurality of processors configured for determining a behavioural profile and/or a behavioural objective; for each of the sequence of information items, performing a sentiment analysis to determine a sentiment score for the respective information item, the sentiment scores for all information items forming a set of sentiment scores for the sequence of information items; determining whether the set of sentiment scores is compatible with the behavioural profile and/or the behavioural objective; responsive to a determination that the set of sentiment scores is sub-optimal in view of the behavioural profile and/or the behavioural objective, resequencing the sequence of information items to optimize the sequence of information in view of the behavioural profile and/or the behavioural objective; and causing the sequence of information items to be presented to the user via a display of a user device.

The proposed system is able to complete different tasks. Each task is defined in terms of a certain amount of information that is delivered and whose usage is required. It will be appreciated that the proposed system is able to operate at the individual level or at the team level.

At the individual level, the system articulates in eight steps detailed below.

At the first step, for each task, the present invention defines the goal to be achieved. Different kinds of goals are considered in the present disclosure. Each goal leads to a different application of the system. The first goal is to improve the mood following the readings of the report. The present invention will use segregation and integration to increase satisfaction and reduce discomfort.

For example, in the presence of "bad" and "good" news, given that the negative impact of bad news is worse than the positive effect of good news, the algorithm will minimize the discomfort by combining the news such that the effects of the negative does not overwhelm the one of the positive. This can be done in terms of a) quantities of news per report, b) sequences of news (first bad and then good news). The output is one report, tailored for the specific user.

The second goal is to sequence the report of information in order to induce an action by the reader, for example, fight inertia. The output is a set of reports, each for each individual user part of the team. The third goal is related to proper usage of information for better individual decision making. The goal of better individual decision making includes two parts. The first part is awareness and usage of all the information avoiding the "information" biases (including but not limited to availability bias, representativeness bias, illusion of control, hot hand fallacy, mean reversion, cognitive inertia, recency and primacy). The second part is allowing decision making not to be biased by "risk-based" biases (e.g., disposition effect, mental accounting, betting with the house money). The fourth goal is related to proper usage of information for better team decision making. In particular, the fourth goal is to provide each member of the team a differently sequenced set of information/documents such that the team-based decision making is optimized. The output is several different reports, one for each user of the team.

At the second step, for each task, the present invention inputs the data/information as a report made of different pieces of information.

Then at the third step, the proposed system clusters the pieces of information together using a textual analysis algorithm that optimizes understandability and minimizes biases. Several documents clustered together represent a block. At the third step, each piece of information is scored as a function of sentiment, congruence, novelty, complexity and length.

Sentiment in the text is identified using a model of textual analysis. Sentiment is based on the tone of the piece of information. An index of tone is created, using positive/negative words. Index of dispersion of tone is built as standard deviation of tone within the piece of information. In some embodiments, sentiment is a function of tone standardized by its dispersion. It is defined between 0 and 100, where 0 is low and 100 is the maximum.

Congruence is defined in terms of logical conclusions and tone. For each piece of information, an index of the degree of congruence of the content with the other pieces of information is created.

Novelty is defined as an index of the difference between the content of the piece of information and the average content of the others.

Complexity defines whether information is simple or complex. Complexity is defined in terms of quantity of numbers given words and numbers, quantity of different words used in the information, whether content is relatively new to the reader, and whether content individuals process large amounts of information on each piece of information/evidence. In particular, whether content is relatively new to the reader is judged according to the number of times the reader has already seen it (based on a counter that keeps track of the number of times the file has been viewed). In one example, whether content individuals process large amounts of information on each piece of information/evidence is judged by whether each information contains more than a certain number of

words. In another example, it is judged by whether the individual is assumed to be familiar or unfamiliar with the information.

The length of a series is based on whether it is a short or long information series. In some embodiments, short information series is classified into 2 to 12 parts of the information sequence, while long information series is categorized if the sequence of information is more than 17 parts.

At the third step, after each piece of information is scored as a function of sentiment, congruence, novelty, complexity, and length, information will be aggregated in blocks. Aggregation is done to make information more understandable and easy to process and to avoid cue competition. Cue competition refers to the fact that, when decision makers process multiple cues differently, some cues will weaken the effect of others, and the impact of irrelevant cues may outweigh the relevant ones thereby causing subjects to use relevant cues less. To avoid cue competition, the pieces of information are positioned in a way that there is no overwhelming piece of information that crowds out the others. Therefore, blocks are constructed so as to have not very dissimilar intensity of sentiment among the pieces of information part of them and to be as little difficult to understand as possible.

In particular, the pieces of information are grouped in blocks using a numerical algorithm that groups pieces of information together by minimizing the product of length and complexity, minimizing the product of novelty, maximizing congruence and minimizing the standard deviation of sentiment across different pieces of information. In some examples, cluster methodology can be applied.

At the fourth step, the proposed system sequences the pieces of information within each block. The pieces of information are sequenced inside the block in order to exploit aggregation/integration bias and reduce segregation bias, increase clarity, and increase attention and awareness.

The fourth step involves increasing proximity of similar content to increase clarity, increase proximity of content with different tone to reduce dissonance

and disposition effect, optimize positioning to avoid attention bias and memory fallacy due to the presence of novel and surprising news.

The sequencing algorithm will optimize along the three dimensions. In particular, the pieces of information are sequenced as a function of the following algorithm. First, a cumulative indicator of sentiment for the set of information is constructed. This is done by sequentially adding pieces of information and calculating the average sentiment for the increasing series of information. Second, within each block a numerical algorithm chooses the sequence of pieces of information that minimizes the standard deviation of the cumulative sentiment indicator, as well as minimizes the occurrence of streaks of sentiments of the same type.

At the fifth step, after the pieces of information are sequenced within each block, refinement is applied to adjust the information/block for primacy and recency. If the block is long, more weight is given to the last pieces of information. If the set of blocks is long, more weight is given to the last presented blocks. If the block is complex, more weight is given to the first pieces of information. If the set of blocks is complex, more weight is given to the first presented blocks. It will be appreciated that weights are provided by either different choice of colour in the provision of the text or in different fonts and style.

At the sixth step, the blocks are sequenced as a function of the following algorithm. First, a cumulative indicator of sentiment for the set of information is constructed. This is done by sequentially adding blocks and calculating the average sentiment for the increasing series of information. Second, a numerical algorithm chooses the sequence of blocks so as to minimize the standard deviation of the cumulative sentiment indicator, and to minimize the occurrence of streaks of sentiments of the same type.

At the seventh step, the optimized report is linked to previous and contemporaneous reports. The system relates evidence and tone to previous or contemporaneous documents and anchors. For each report, the present invention defines a set of closely-related reports in the previous N days. Reports are defined

as closely-related in terms of text/content and share a common theme. Moreover, the sentiment, congruence, novelty, complexity and length of the previous reports are recorded. In addition, the current report is compared to the average/most relevant of the previous reports. The goal is to process the current optimized report in light of the information contained in the previous reports.

The eighth step relates to ex-post assessment. When the feedback on the outcome of the decision is available, an assessment of the link between the characteristics of the report and the quality of the decision is made to provide evidence on biases or user/team.

In the case of team level decision, the above eight-step algorithm needs to be modified in order to amplify the discrepancy in the way the report is provided to each individual part of the team. At the team level, higher herding requires more mixing up of reports tones for the different users. In particular, the previous eight-step algorithm will be modified in two ways detailed below.

The first way is to store in the report the degree of consensus reach for each decision based on each report. The second way is to maximize the distance in the sentiment-exposure profile for each participant. In particular, the proposed system proceeds as follows. For each individual, the numerical algorithm that chooses the sequence of pieces of information within the blocks and the sequencing of the block themselves is modified in order to maximize the standard deviation of the cumulative sentiment indicator for each individual, maximize the occurrence of streaks of sentiments of the same type, for each individual, and increase the difference between the maximized cumulative sentiment indicator across the individuals.

Figure 2 is a block diagram showing an exemplary computer device 300, in which embodiments of the invention may be practiced. The computer device 300 may be a mobile computer device such as a smart phone, a wearable device, a palm-top computer, and multimedia Internet enabled cellular telephones, an on-board computing system or any other computing system, a mobile device such as an iPhone™ manufactured by Apple™, Inc or one manufactured by LG™, HTC™ and Samsung™, for example, or other device.

As shown, the mobile computer device 300 includes the following components in electronic communication via a bus 306:

- (a) a display 302;
- (b) non-volatile (non-transitory) memory 304;
- (c) random access memory ("RAM") 308;
- (d) N processing components 310;
- (e) a transceiver component 312 that includes N transceivers; and
- (f) user controls 314.

Although the components depicted in Figure 2 represent physical components, Figure 2 is not intended to be a hardware diagram. Thus, many of the components depicted in Figure 2 may be realized by common constructs or distributed among additional physical components. Moreover, it is certainly contemplated that other existing and yet-to-be developed physical components and architectures may be utilized to implement the functional components described with reference to Figure 2.

The display 302 generally operates to provide a presentation of content to a user, and may be realized by any of a variety of displays (e.g., CRT, LCD, HDMI, micro-projector and OLED displays).

In general, the non-volatile data storage 304 (also referred to as non-volatile memory) functions to store (e.g., persistently store) data and executable code. The system architecture may be implemented in memory 304, or by instructions stored in memory 304.

In some embodiments for example, the non-volatile memory 304 includes bootloader code, modem software, operating system code, file system code, and code to facilitate the implementation components, well known to those of ordinary skill in the art, which are not depicted nor described for simplicity.

In many implementations, the non-volatile memory 304 is realized by flash memory (e.g., NAND or ONENAND memory), but it is certainly contemplated that other memory types may be utilized as well. Although it may be possible

to execute the code from the non-volatile memory 304, the executable code in the non-volatile memory 304 is typically loaded into RAM 308 and executed by one or more of the N processing components 310.

The N processing components 310 in connection with RAM 308 generally operate to execute the instructions stored in non-volatile memory 304. As one of ordinary skill in the art will appreciate, the N processing components 310 may include a video processor, modem processor, DSP, graphics processing unit (GPU), and other processing components.

The transceiver component 312 includes N transceiver chains, which may be used for communicating with external devices via wireless networks. Each of the N transceiver chains may represent a transceiver associated with a particular communication scheme. For example, each transceiver may correspond to protocols that are specific to local area networks, cellular networks (e.g., a CDMA network, a GPRS network, a UMTS networks), and other types of communication networks.

The system 300 of Figure 2 may be connected to any appliance 418, such as one or more cameras mounted to the vehicle, a speedometer, a weather service for updating local context, or an external database from which context can be acquired.

It should be recognized that Figure 2 is merely exemplary and in one or more exemplary embodiments, the functions described herein may be implemented in hardware, software, firmware, or any combination thereof. If implemented in software, the functions may be stored on or transmitted over as one or more instructions or code encoded on a non-transitory computer-readable medium 304. Non-transitory computer-readable medium 304 includes both computer storage medium and communication medium including any medium that facilitates transfer of a computer program from one place to another. A storage medium may be any available medium that can be accessed by a computer.

It will be appreciated that many further modifications and permutations of various aspects of the described embodiments are possible. Accordingly, the

described aspects are intended to embrace all such alterations, modifications, and variations that fall within the spirit and scope of the appended claims.

Throughout this specification and the claims which follow, unless the context requires otherwise, the word "comprise", and variations such as "comprises" and "comprising", will be understood to imply the inclusion of a stated integer or step or group of integers or steps but not the exclusion of any other integer or step or group of integers or steps.

The reference in this specification to any prior publication (or information derived from it), or to any matter which is known, is not, and should not be taken as an acknowledgment or admission or any form of suggestion that that prior publication (or information derived from it) or known matter forms part of the common general knowledge in the field of endeavour to which this specification relates.

Claims:

1. A computer-implemented method of presenting a sequence of information items to a user, the method comprising:
 - determining a behavioural profile and/or a behavioural objective;
 - for each of the sequence of information items, performing a sentiment analysis to determine a sentiment score for the respective information item, the sentiment scores for all information items forming a set of sentiment scores for the sequence of information items;
 - determining whether the set of sentiment scores is compatible with the behavioural profile and/or the behavioural objective;
 - responsive to a determination that the set of sentiment scores is sub-optimal in view of the behavioural profile and/or the behavioural objective, resequencing the sequence of information items to optimize the sequence of information in view of the behavioural profile and/or the behavioural objective; and
 - causing the sequence of information items to be presented to the user via a display of a user device.
2. A method according to claim 1, wherein the behavioural objective is a mood of the user, a risk-taking capacity of the user, a change in attention bias of the user, or a difference in attitude of the user to one or more other users.
3. A method according to claim 1 or claim 2, wherein said resequencing comprises modifying one or more of: the order of the information items; temporal spacing between successive information items; and sub-groupings of the information items within the sequence.
4. A method according to claim 3, wherein said resequencing comprises adjusting temporal spacing between successive information items within a user-allocated viewing time period.

5. A method according to any one of the preceding claims, wherein the set of sentiment scores is determined at least partly based on metadata associated with the information items.
6. A method according to any one of the preceding claims, wherein resequencing comprises reordering/resequencing the information items so that the set of sentiment scores is compatible with the behavioural profile and/or the behavioural objective.
7. A system of presenting a sequence of information items to a user, the system comprising a plurality of processors configured for:
 - determining a behavioural profile and/or a behavioural objective;
 - for each of the sequence of information items, performing a sentiment analysis to determine a sentiment score for the respective information item, the sentiment scores for all information items forming a set of sentiment scores for the sequence of information items;
 - determining whether the set of sentiment scores is compatible with the behavioural profile and/or the behavioural objective;
 - responsive to a determination that the set of sentiment scores is sub-optimal in view of the behavioural profile and/or the behavioural objective, resequencing the sequence of information items to optimize the sequence of information in view of the behavioural profile and/or the behavioural objective; and
 - causing the sequence of information items to be presented to the user via a display of a user device.
8. A system according to claim 7, wherein the behavioural objective is a mood of the user, a risk-taking capacity of the user, a change in attention bias of the user, or a difference in attitude of the user to one or more other users.
9. A system according to claim 7 or claim 8, wherein said resequencing comprises modifying one or more of: the order of the information items; temporal spacing between successive information items; and sub-groupings of the information items within the sequence.

- 10.A system according to claim 9, wherein said resequencing comprises adjusting temporal spacing between successive information items within a user-allocated viewing time period.
- 11.A system according to claims 7 to 10, wherein the set of sentiment scores is determined at least partly based on metadata associated with the information items.
- 12.A system according to claims 7 to 11, wherein resequencing comprises resequencing the information items so that the set of sentiment scores is compatible with the behavioural profile and/or the behavioural objective.