

## **ABSTRACT**

Systems and methods for improved management of sentiments over conventional approaches are disclosed. Supervised approach is used to augment the rule-based approach for classification. Initially, sentiment evaluation is performed by the system using a rule based approach and an interface is provided to the user to give feedback on the correctness of evaluated sentiment. This feedback is used by the sentiment evaluation system to update the set of rule-based and also apply the supervised approach to train the classifier for evaluating complex posts.

REF FIG: 1

What is claimed is:

1. A computer-implemented method executed by one or more computing devices for managing sentiments, the method comprising:
  - receiving, using the one or more computing devices, one or more evaluated posts as an input item to be managed, wherein each of the one or more evaluated posts is categorized into a first classification type;
  - identifying by a classifier, using the one or more computing devices, each of the one or more evaluated posts in a multidimensional feature space, wherein each dimension represents a unique feature corresponding to the one or more evaluated posts;
  - receiving, using the one or more computing devices, an indication from at least one user on each of the one or more identified evaluated posts, as an output item to be managed, wherein the indication from the user represents a second classification type; and
  - performing by the classifier, using the one or more computing devices, a neighborhood operation for each of the one or more identified evaluated posts, based on the second classification type.
2. The computer-implemented method in accordance with claim 1, further comprising the step of selecting a preconfigured subset of features in the multidimensional feature space.
3. The computer-implemented method in accordance with claim 1, further comprising the step of performing the neighborhood operation by adjusting a preconfigured pattern space around each of the one or more evaluated posts identified in the multidimensional feature space.
4. The computer-implemented method in accordance with claim 1, further comprising the step of performing the neighborhood operation by re-classifying the first classification type for each of the one or more posts in a pattern space.
5. The computer-implemented method in accordance with claim 3, further comprising the step of performing the neighborhood operation by re-classifying the first classification type for each of the one or more posts in the pattern space.
6. The computer-implemented method in accordance with claim 5, further comprising the step of subjecting each of the one or more re-classified posts as training data sets for subsequent classification, wherein the subsequent classification corresponds to the first classification type.

7. The computer-implemented method in accordance with claim 1, wherein the step of performing the neighborhood operation comprises of adapting at least one pre-configured pattern space.
8. A system for managing sentiments, the system comprising:
  - an input terminal for receiving one or more evaluated posts as an input item to be managed, wherein each of the one or more evaluated posts is categorized into a first classification type; and
  - a computing system communicating with the input terminal comprising, the computing system having a processor and a memory, the memory capable of storing software components for execution by the processor, the software components comprising:
    - an identification module operative to identify each of the one or more evaluated posts in a multidimensional feature space, wherein each dimension represents a unique feature corresponding to the one or more evaluated posts;
    - an indication module for receiving an indication from at least one user on each of the one or more identified evaluated posts, as an output item to be managed, wherein the indication from the user represents a second classification type; and
    - an operational module for performing a neighborhood operation for each of the one or more evaluated based on the second classification type.
9. The system in accordance with claim 8, wherein the identification module is further operative to select a preconfigured subset of features in the multidimensional feature space.
10. The system in accordance with claim 8, wherein the operational module is further operative to adjust a preconfigured pattern space around each of the one or more evaluated posts identified in the multidimensional feature space.
11. The system in accordance with claim 8, wherein the operational module is further operative to perform the step of re-classifying the first classification type for each of the one or more posts in the pattern space.
12. The system in accordance with claim 10, wherein the neighborhood operation module is further operative to re-classify the first classification type for each of the one or more posts in the pattern space.

13. The system in accordance with claim 12, further operative to subject each of the one or more re-classified posts as training data sets for subsequent classification, wherein the subsequent classification corresponds to the first classification type.
14. The system in accordance with claim 8, wherein the neighborhood operational module is further operative to adapt at least one pre-configured pattern space.
15. A computer readable medium having a set of instructions for execution on a computing device, the set of instructions comprising:
- a receiving routine operative for receiving one or more evaluated posts as an input item to be managed, wherein each of the one or more evaluated posts is categorized into a first classification type;
  - an identification routine operative to identify each of the one or more evaluated posts in a multidimensional feature space, wherein each dimension represents a unique feature corresponding to the one or more evaluated posts;
  - an indication routine operative for receiving an indication from at least one user on each of the one or more identified evaluated posts, as an output item to be managed, wherein the indication from the user represents a second classification type; and
  - an operational routine operative for performing a neighborhood operation for each of the one or more evaluated based on the second classification type.

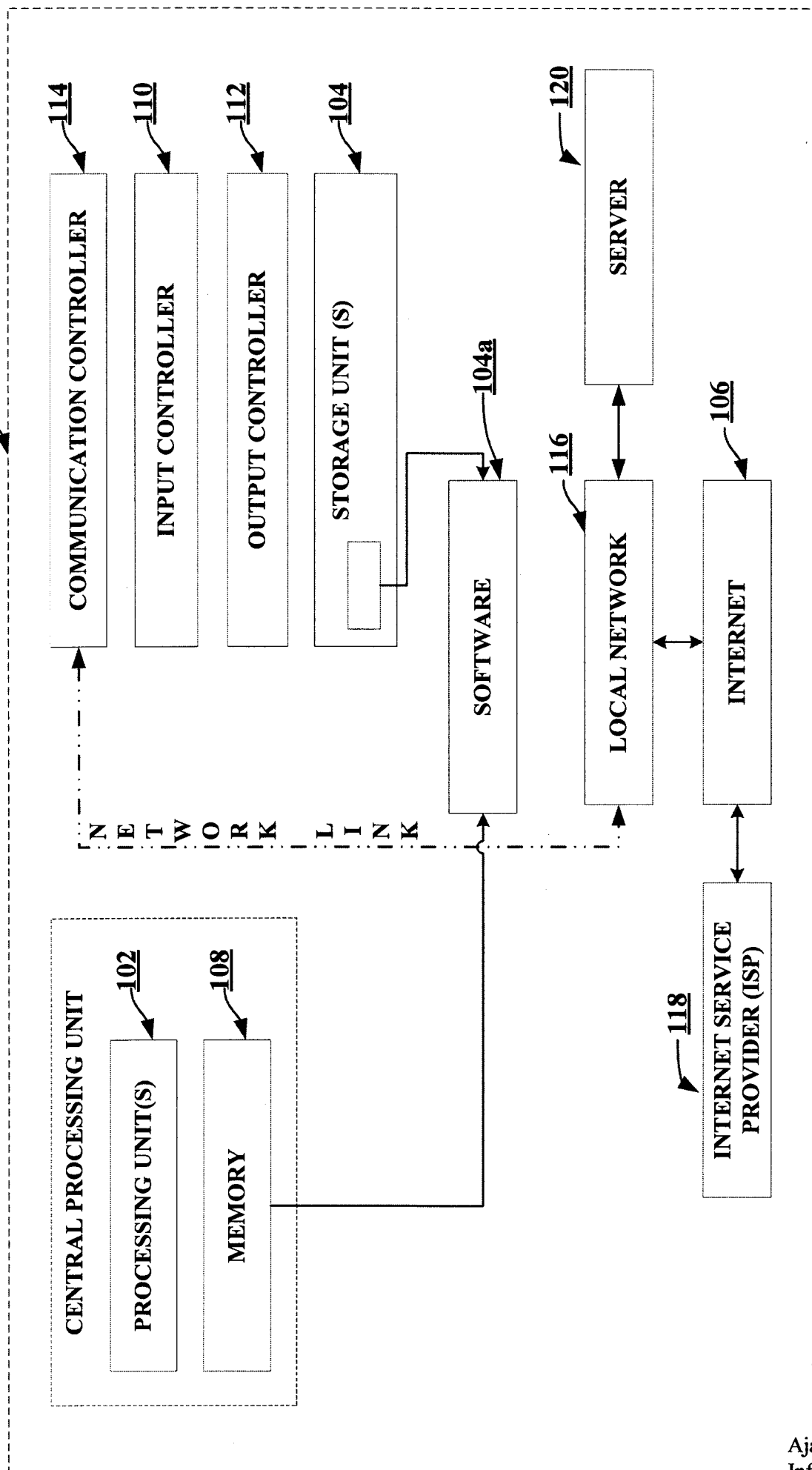
Dated this 18th day of March, 2013



Ajay Kumar Sarkar

Infosys Limited

Patent Agent No 1652

FIGURE 1  
(PRIOR-ART)

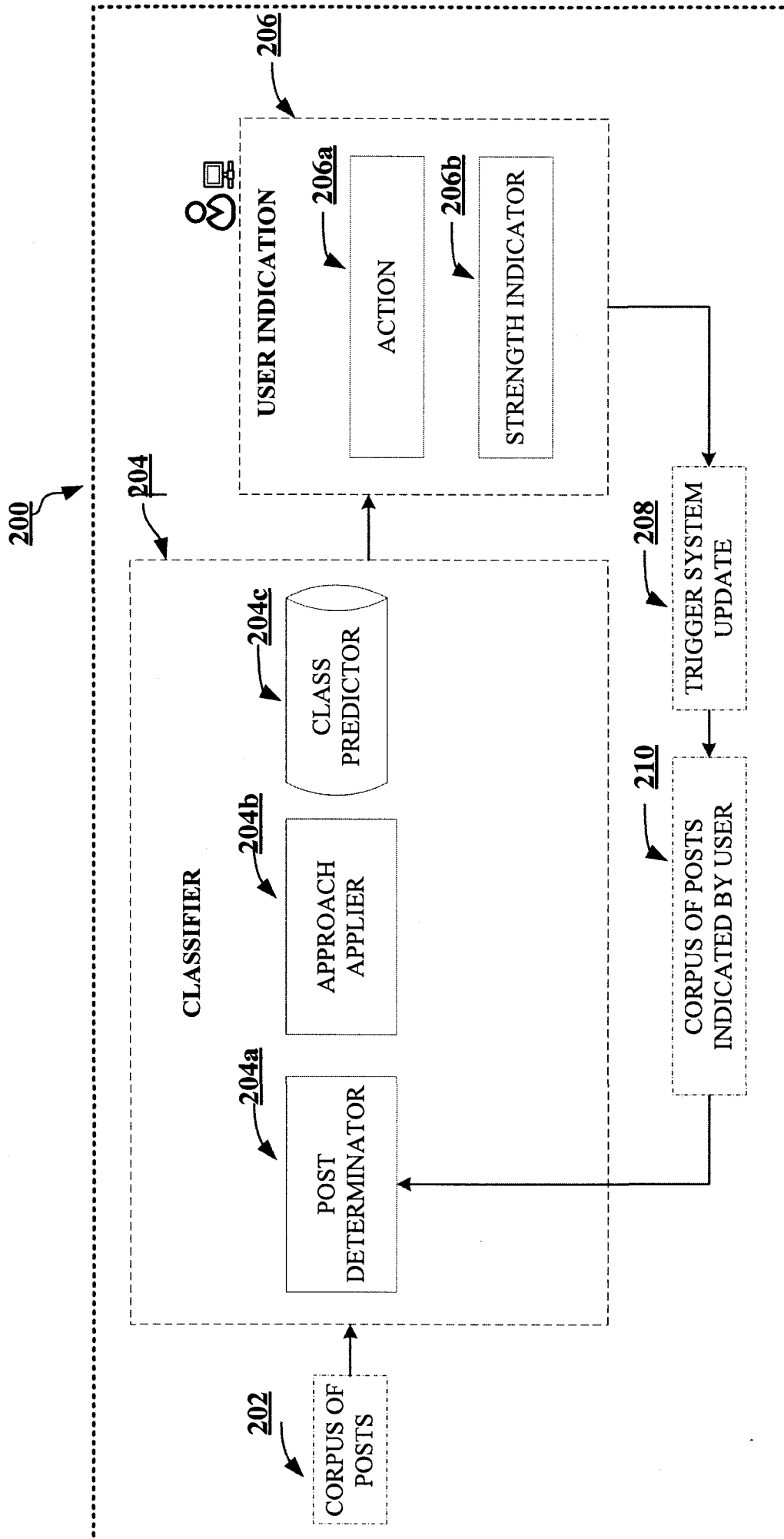



FIGURE 2

  
Ajay Kumar Sarkar  
Infosys Limited  
Agent No. : 1652

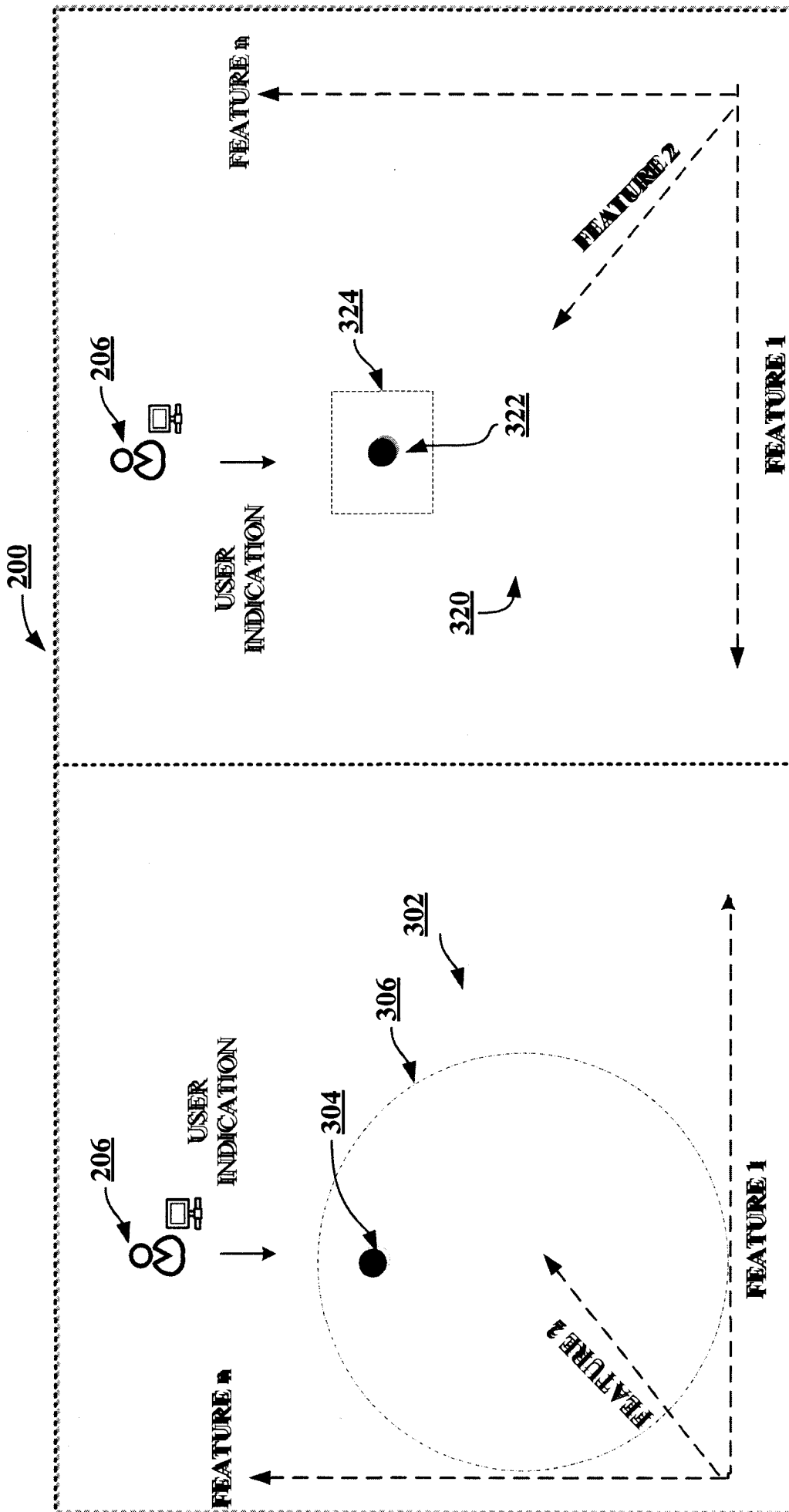


FIGURE 3

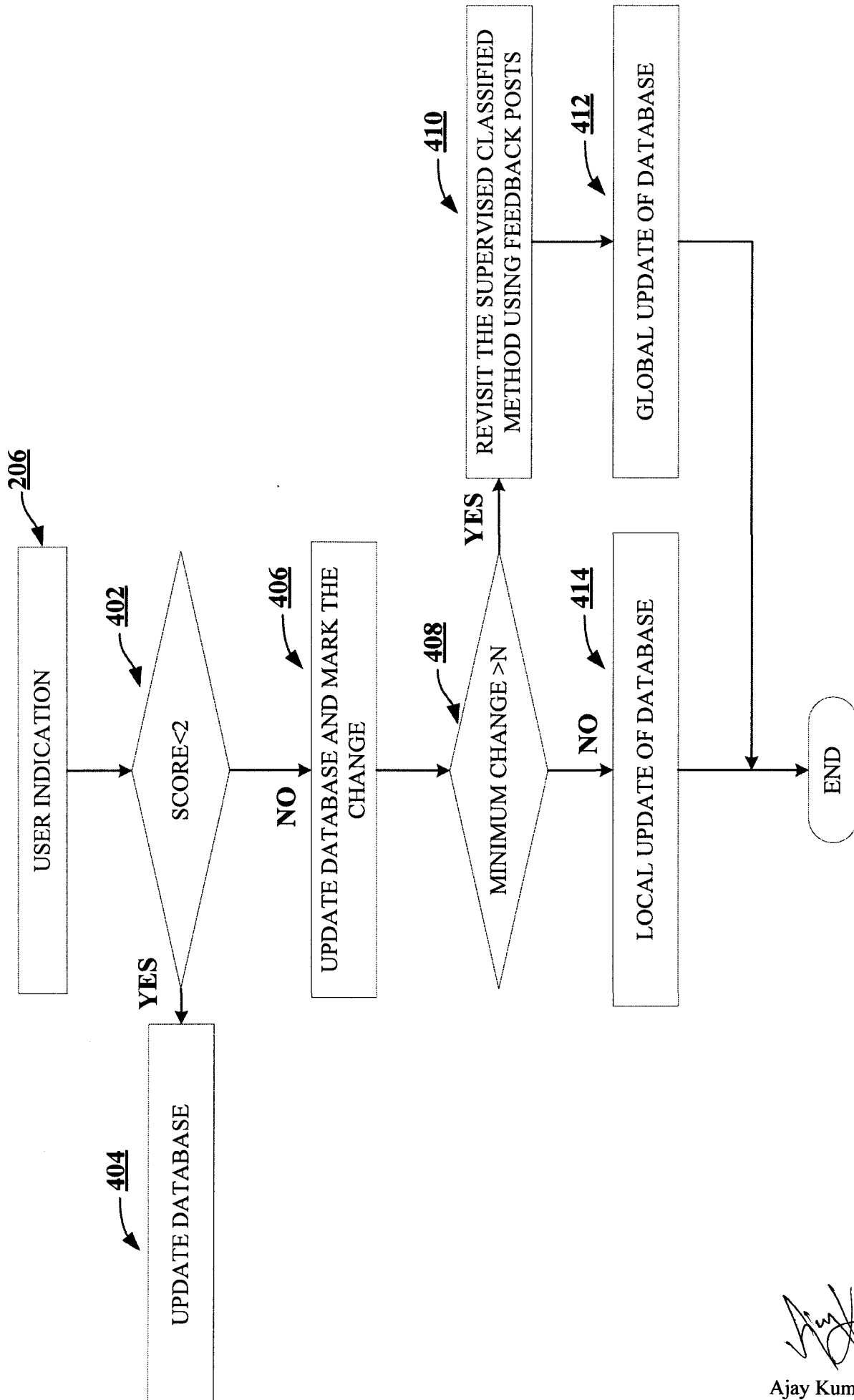



FIGURE 4

  
Ajay Kumar Sarkar  
Infosys Limited  
Agent No. : 1652



# **SYSTEM FOR MANAGEMENT OF SENTIMENTS AND METHODS THEREOF**

## **FIELD OF THE INVENTION**

[0001] The present disclosure relates generally to the field of classification of content, and, more particularly, to systems and methods for sentiment classification and management.

## **BACKGROUND OF THE INVENTION**

[0002] A content classification system (hereinafter may be referred to as 'classification system'), in general, might be used to filter content from a large corpus to identify content in context of some information. The classification could also be used to flag the content in terms of belonging to a distinct class. For example, a content classification system could be used to capture sentiments from posts relating to one or more products or services, from various forums, including social platforms. The term, 'post', as used herein, means an amalgamation of sentences. The term, social platform, as used herein, means web-based services that allow members to construct a public or a semi-public profile within the boundary of that particular community. Classification of these posts could be used to evaluate a marketing strategy for a product or service. While a post is used as an example of content, it should be understood that any artifact with an amalgamation of sentences might be equally applicable.

[0003] A sentiment, in general, refers to a view or an opinion towards a situation or an event. Various methods have been proposed for evaluating sentiment from text. Conventionally, evaluation of sentiment from text within the posts has been performed using a rule based approach. Ruled-based approach applies a set of rules to pre-process text in order to aid classification by techniques such as, part-of-speech (hereinafter may be referred to as 'POS') tagging and keyword phrases to tag sentiment into distinct classes. Rule-based approach needs to be exhaustive and deriving sentiment from free form text using incomplete rules can result in erroneous evaluation. Evolving rule-base over time can be difficult as changing POS tagging rules may not be feasible for a corpus containing large number of posts. Further, predefined static rules might be infeasible to cover all classes of possible text examples while conceiving the system.

[0004] Another approach has been to use supervised learning methods where several features from a corpus of posts are derived to train supervised classifiers. Training might involve providing the system with a corpus and classifications for each post in the training corpus. The system then operates a training process wherein patterns are discovered in the training corpus between the classifications of various posts. Supervised approach can also be limiting for system evolution as it requires re-training the entire system when new evidence is provided.

[0005] While these systems are useful, it is difficult to use either of the two approaches when the number of posts is not sufficient to build an exhaustive set of rules or appropriately train a classifier. Further, sentiment evaluation might be subjective to a user and therefore the same post can convey different information to another user. Hence, there is a considerable need in the art for a more sophisticated system capable of managing sentiments based on multiple inputs into multidimensional categories.

[0006] Disclosed herein are improved systems and methods for sentiment management which is adaptive to a user.

### **SUMMARY OF THE DISCLOSURE**

[0007] Aspects of the disclosure relate to systems and methods for managing sentiments by making it adaptive to a user.

[0008] It is therefore one object of the present disclosure to provide systems and methods to use multiple approaches to derive sentiments and receive indications from at least one user pertaining to at least one post from a corpus of posts.

[0009] It is another object of the present disclosure to perform a neighborhood operation for posts with user indications.

[0010] It is yet another object of the present disclosure to enable a local and global update of the system for posts with user indications, for same or similar posts.

[0011] The above as well as additional aspects and advantages of the disclosure will become apparent in the following detailed written description

### **BRIEF DESCRIPTION OF THE DRAWINGS**

[0012] The aspects of the disclosure will be better understood with the accompanying drawings.

[0013] **FIGURE 1 (PRIOR ART)** is a block diagram of a computing device to which the present disclosure may be applied.

[0014] **FIGURE 2** is a high-level block diagram illustrating a system **200** in accordance with an embodiment of the present disclosure.

[0015] **FIGURE 3** is an illustration of a representative post from a corpus of posts **210** for which an update is triggered **208** in the system **200**

[0016] **FIGURE 4** illustrates a method to update the system based on user indications, in accordance with an embodiment of the present disclosure.

[0017] While systems and methods are described herein by way of example and embodiments, those skilled in the art recognize that systems and methods disclosed herein are not limited to the embodiments or drawings described. It should be understood that the drawings and description are not intended to be limiting to the particular form disclosed. Rather, the intention is to cover all modifications, equivalents and alternatives falling within the spirit and scope of the appended claims. Any headings used herein are for organizational purposes only and are not meant to limit the scope of the description or the claims. As used herein, the word “may” is used in a permissive sense (i.e., meaning having the potential to) rather than the mandatory sense (i.e., meaning must). Similarly, the words “include”, “including”, and “includes” mean including, but not limited to.

#### **DETAILED DESCRIPTION**

[0018] Disclosed embodiments provide computer-implemented methods, systems, and computer-readable media for managing sentiments by making it adaptive to a user. A user refers to at least one user who is using the sentiment management system. The embodiments described herein are related to management of posts to derive sentiments. While the particular embodiments described herein may illustrate the disclosure in a particular domain, the broad principles behind these embodiments could be applied in other fields of endeavor. To facilitate a clear understanding of the present disclosure, illustrative examples are provided herein which describe certain aspects of the disclosure. However, it is to be appreciated that these illustrations are not meant to limit the scope of the disclosure, and are provided herein to illustrate certain concepts associated with the disclosure.

[0019] It is also to be understood that the present disclosure may be implemented in various forms of hardware, software, firmware, special purpose processors, or a combination thereof. Preferably, the present disclosure is implemented in software as a program tangibly embodied on a program storage device. The program may be uploaded to, and executed by, a machine comprising any suitable architecture.

[0020] **FIG. 1 (PRIOR-ART)** is a block diagram of a computing device **100** to which the present disclosure may be applied according to an embodiment of the present disclosure. The system includes at least one processor **102**, designed to process instructions, for example computer readable instructions (i.e., code) stored on a storage device **104**. By processing instructions, processing device **102** may perform the steps and functions disclosed herein. Storage device **104** may be any type of storage device, for example, but not limited to an optical storage device, a magnetic storage device, a solid state storage device and a non-transitory storage device. The storage device **104** may contain software **104a** which is a set of instructions (i.e. code). Alternatively, instructions may be stored in one or more remote storage devices, for example storage devices accessed over a network or the internet **106**. The computing device also includes an operating system and microinstruction code. The various processes and functions described herein may either be part of the microinstruction code or part of the program (or combination thereof) which is executed via the operating system. Computing device **100** additionally may have memory **108**, an input controller **110**, and an output controller **112** and communication controller **114**. A bus (not shown) may operatively couple components of computing device **100**, including processor **102**, memory **108**, storage device **104**, input controller **110** output controller **112**, and any other devices (e.g., network controllers, sound controllers, etc.). Output controller **110** may be operatively coupled (e.g., via a wired or wireless connection) to a display device (e.g., a monitor, television, mobile device screen, touch-display, etc.) in such a fashion that output controller **110** can transform the display on display device (e.g., in response to modules executed). Input controller **108** may be operatively coupled (e.g., via a wired or wireless connection) to input device (e.g., mouse, keyboard, touch-pad, scroll-ball, touch-display, etc.) in such a fashion that input can be received from a user. The communication controller **114** is coupled to a bus (not shown) and provides a two-way coupling through a network link to the internet **106** that is connected to a local network **116** and operated by an internet service provider (hereinafter referred to as 'ISP') **118** which provides data communication services to the internet. Network link typically provides data communication through one or more networks to other data devices. For example, network link may provide a connection through local network **116** to a host computer, to data equipment operated by an ISP **118**. A server **120** may

transmit a requested code for an application through internet 106, ISP 118, local network 116 and communication controller 114. Of course, FIG. 1 illustrates computing device 100 with all components as separate devices for ease of identification only. Each of the components may be separate devices (e.g., a personal computer connected by wires to a monitor and mouse), may be integrated in a single device (e.g., a mobile device with a touch-display, such as a smartphone or a tablet), or any combination of devices (e.g., a computing device operatively coupled to a touch-screen display device, a plurality of computing devices attached to a single display device and input device, etc.). Computing device 100 may be one or more servers, for example a farm of networked servers, a clustered server environment, or a cloud network of computing devices.

[0021] Usage of systems and methods described herein provide for improved management of sentiments over conventional approaches. Supervised approach is used to augment the rule based approach for classification. Initially, sentiment evaluation is performed by the system using a rule based approach and an interface is provided to the user to give feedback on the correctness of evaluated sentiment. This feedback is used by the sentiment evaluation system to update the set of rule base and also apply the supervised approach to train the classifier for evaluating complex posts. Posts can be classified into discrete classes for a given indication from a user. The term, 'post', as used herein, refers to an amalgamation of sentences. The terms, post and corpus of posts are interchangeably herein. Posts are input items for a classifier. For example, a post may include, but is not limited to, text, an image, a combination of images and text, or other organizations of content. The term 'classifier' refers to a system or code for mapping posts to discrete classes. The term 'classifier', as used herein, may be used interchangeably with 'sentiment evaluation system'. Classifier may use one or more approaches to classify sentiments evaluated from posts. For example, in a social platform, there might be a collection of posts relating to a product review with references to relating products.

[0022] In many cases, a corpus might be divided into classes one way for one indication and would be divided into those classes differently for an indication from a second user. The classifier, based on the type of approach employed, might classify a post as belonging to a negative class. There might be indications by several hundreds of users pertaining to the product review. From the indication of a

first user, the post may be classified as belonging to a positive class and from the indication of a second user; the same post may be classified as belonging to a neutral class. Conventional approaches would only recognize the classification made by the classifier. However, in embodiments of the present disclosure as described herein, classifications made by the classifier are adapted to a user and further applied for future posts pertaining to same or similar contextual information.

[0023] **FIG. 2** is a high-level block diagram illustrating a system **200** in accordance with an embodiment of the present disclosure. When a post or a corpus of posts **202** is presented as an input item, system **200** operates a classifier **204** which processes posts to derive sentiments. Sentiment analysis can be performed for a product, product line and brand against time, region and population. For example, social platforms like blogs, Twitter, Facebook and the like provide useful channels to capture user sentiment early in comparison to traditional mediums like pre-launch consumer survey or the first set of sales figures. The posts on these social platforms may be analyzed for a given product or a product line and statistics on various sentiments may be generated for further analysis. According to an embodiment of the present disclosure, sentiments are derived by extracting linguistic components from each of the posts by generating features from input item to be considered for sentiment rating or evaluation. The term, feature, as used herein, refers to a unique attribute that could affect the derivation of sentiments from the post. Feature generation is primarily performed for efficiency reasons in order to increase the consistency of deriving sentiments. A post discriminator **204a** generates from the input item, one or more features from post to be considered. Features might be domain specific, for example, in the case of English text classification; a relevant feature might be the presence of a unique word within a document, regardless of position. Based on the features and the linguistic extraction, a post determinator **204a** determines the nature of a post within the corpus based on a complexity score for that post. According to an embodiment of the disclosure, a post may be determined as simple or complex. The term, 'simple post', as used herein, refers to smaller statements with few POS and a simple structure. The term, 'complex post', as used herein, refers to multiple sentences. For example, the post determinator **204a** may be pre-configured to calculate a complexity score based on the linguistics, for example, but not limited to, the

presence of pronouns, words and conjunctions in a post so as to classify a post from very simple to very complex.

LINGUISTIC EXTRACTION			Score	Post Determination
Pronouns (denoted as 'a')	Words (denoted as 'b')	Conjunctions		
$\leq a$	$\leq b$	NA	0	very simple
$> a$	$\leq b$	NA	1	simple
NA	$< b$	$> c$	2	complex
NA	$\geq b$	NA	3	very complex

**Table I**

[0024] Once the post determination is completed, classifier operates an approach applier **204b** employs at least one approach to a post so as to determine a first classification type for that post, using a class predictor **204c**. The approach applier **204b** can preferably be pre-configured to apply a rule-based approach for posts classified as very simple and simple posts. Complex posts may be applied with either rule-based approach or supervised approach or both approaches in a sequence. The implementation of either one or both approaches generates a discrete predicted class. The class predictor **204c** classifies simple posts into discrete classes, for example, positive, negative or neutral sentiments. Complex posts may be classified as positive, negative, neutral and non-obvious. In other variations, the degree of the first classification type may be possible to yield more than four classes, for example, strongly positive and slightly positive. Posts with non-obvious sentiments can be routed through the approach applier **204b** once again to employ the supervised approach for re-classifying using the class predictor **204c**. Given the sentiment evaluated for posts by the class predictor **204c**, a user **206** can either accept or reject the sentiment value. An action **206a** by a user can trigger an update **208** in the sentiment evaluation system which modifies the system backend and also alters the way sentiment is evaluated for new posts in future. For example, a current post has been marked as belonging to a discrete class identified as a positive sentiment, then the reject action results in sentiment being modified to negative, and, vise-versa. A system update is triggered accordingly for future sentiment evaluation. Strength Indicator **206b** captures the second classification type of the sentiment classes derived by the classifier **204**.

The impact of indication on the classifier **204** is determined by a second classification type **206b**. While taking an action in relation to a post, a user can assign confidence values for the action **206a** taken in relation to the discrete class identified for a post. According to an embodiment of this disclosure, confidence values could have variations which may include gradations indicating a discrete or predicted class as 'very satisfied', satisfied, 'dissatisfied' or 'very dissatisfied'. If the indication 'very satisfied' or 'very dissatisfied' then a larger number of posts is likely to be impacted whereas the effect of 'satisfied' or 'dissatisfied' is likely to have a marginal impact. For example, there may be posts about a product launch. A requirement may be to derive sentiments of social platform users' opinion for the launched product. Classifier **204** can take these posts and derive sentiments. The classifier **204** can work through the system to capture indications from other users. The systems and methods of the disclosure can be used to employ supervised approach for continued learning. After training, the data classification system can be used to predict classification of an unlabeled instance or a non-obvious instance.

[0025] **FIG. 3** is an illustration of a representative post from a corpus of posts **210** for which an update is triggered **208** in the system **200** in accordance with the method illustrated in **FIG. 4**. Based on the user indication **206** recorded by the system **200**, an update is triggered in the system **208** for the corpus of posts **210** for which a user indication **206** has been received. Classifier **204** accepts a corpus of posts indicated by a user **210** comprising of at least one post. If the post determinator **206a** determines the post as 'simple' or 'very simple' **402**, an update is made to at least database to record the change **404**. The sentiment of the post or the predicted class **206b** is updated on the system. Preferably, no change occurs in the way sentiment class is predicted for future posts. A new rule can be added to the rule-based approach to evaluate posts corresponding to feature values of each of these posts. On the other hand, if the post determinator **206a** determines the post as 'complex' or 'very complex', the update is made to at least database and the change is marked. **406**. According to an embodiment of the disclosure, if sufficient number of indications (denoted by N) is received **408**, the classifier **204** operates the approach applier **204b** to employ supervised approach **410** to train the system for future classification of same or similar posts. The supervised classifier is trained or retrained for complex posts with a particular complexity score.



[0026] Every post in the corpus can be observed as a point in a multidimensional orthogonal numerical feature space. Each dimension corresponds to a feature used for training the classifier. Example include, but is not limited to, word frequency, POS frequency, co-occurrence frequency ordering, website impact metric, user impact metric, word length of post, word distance adjective mapping, pronoun mapping and resolving conjunction. A subset of these features (feature 1, feature 2 and feature n) is used to describe each post in the feature space for a given domain. After a post is distilled into a set of enumerated features, classifier **204** performs a neighborhood operation in the representative feature spaces **302** and **320**. Position of a post **304**, **322** (hereinafter may be referred to as a 'feature vector') along with user indication on sentiment is identified in the respective feature space **302** and **320**. The term, neighborhood operation, as used herein, refers to range of same or similar posts which are updated for previously identified discrete class, in-line with the strength indication **206b**. Generally, the more the classifier **204** is taught the more it is trained in order to improve the consistency. The neighborhood operation may be applied globally **306**, **412** or locally **324**, **414**. A global neighborhood operation **306**, **412** essentially extends the user indication all existing sentiment classifications derived from a previously same or similar post. No change in classifier is made for simple posts. For example, the discrete classes may extend to a post relating to a subject such as a company, competitor, or the like. A local neighborhood operation **324**, **414** on the other hand makes changes to closely related posts only. The neighborhood operation can be performed by adjusting a preconfigured pattern space in n-dimensions around each of the one or more evaluated posts identified in the multidimensional feature space. Pre-configured pattern spaces include, for example, spheroid, ellipsoid, rectangular, or the like. The term pattern space, as used herein, is used to identify the local neighborhood of the post. The size of neighborhood may be controlled using these pre-configured pattern space with predefined values to determine the extent of feature space based on the strength indications **206b**. The predefined pattern spaces are applied dynamically during system update **208** based on the subset of enumerated features, identifying a feature vector.

[0027] 
$$R = \frac{K_i}{2n} \sum_{i=1}^n |\max(f_n) - \min(f_n)| \quad \text{(Equation 1 for Spheroid)}$$

$$\begin{aligned}
A &= A_1 = K_1 | \max(f_1) - \min(f_1) | \\
B &= A_2 = K_2 | \max(f_2) - \min(f_2) | \\
&\dots \\
A_n &= K_n | \max(f_n) - \min(f_n) |
\end{aligned}$$

[0028]

[0029] General equation of a centered ellipsoid -  $\sum_{i=1}^n \frac{x_i^2}{A_i^2} = 1$  (Equation 2 for Ellipsoid)

$$\begin{aligned}
L1 &= K_1 | \max(f_1) - \min(f_1) | \\
L2 &= K_2 | \max(f_2) - \min(f_2) | \\
&\dots \\
Ln &= K_n | \max(f_n) - \min(f_n) |
\end{aligned}$$

[0030] (Equation 3 for Rectangle)

[0031] Where:

[0032] K= Value of integer  $K_i$  is pre-configured and  $K_i = [0-1]$ . K is large when strength indication has a large weight and vice-versa.

[0033] Using the formula of Equations 1, 2 and 3 for a spheroid pattern space, ellipsoid pattern space and a rectangular pattern space respectively, a neighborhood operation can be defined for the given post.

[0034] Various modifications and variations of the present disclosure will be apparent to those skilled in the art without departing from the scope and spirit of the disclosure. Although the disclosure has been described in connection with specific preferred embodiments, it should be understood that the disclosure as claimed should not be unduly limited to such specific embodiments. Indeed, various modifications of the described modes for carrying out the disclosure which are obvious to those skilled in the art are intended to be within the scope of the claims.