



MINISTERO DELLO SVILUPPO ECONOMICO  
DIREZIONE GENERALE PER LA LOTTA ALLA CONTRAFFAZIONE  
UFFICIO ITALIANO BREVETTI E MARCHI

|                                     |                        |
|-------------------------------------|------------------------|
| <b>DOMANDA DI INVENZIONE NUMERO</b> | <b>102018000008086</b> |
| <b>Data Deposito</b>                | <b>14/08/2018</b>      |
| <b>Data Pubblicazione</b>           | <b>14/02/2020</b>      |

Classifiche IPC

| <b>Sezione</b> | <b>Classe</b> | <b>Sottoclasse</b> | <b>Gruppo</b> | <b>Sottogruppo</b> |
|----------------|---------------|--------------------|---------------|--------------------|
| G              | 06            | Q                  | 10            | 10                 |

| <b>Sezione</b> | <b>Classe</b> | <b>Sottoclasse</b> | <b>Gruppo</b> | <b>Sottogruppo</b> |
|----------------|---------------|--------------------|---------------|--------------------|
| G              | 06            | Q                  | 30            | 02                 |

Titolo

|                                                         |
|---------------------------------------------------------|
| Elaborazione di messaggi scambiati su un social network |
|---------------------------------------------------------|

## **DESCRIZIONE**

Della Domanda di Brevetto per Invenzione Industriale dal titolo:

**“ELABORAZIONE DI MESSAGGI SCAMBIATI SU UN SOCIAL NETWORK”**

a nome : 1) *Telecom Italia S.p.A.*

2) *POLITECNICO DI TORINO*

con sede in : 1) MILANO (Italia)

2) TORINO (ITALIA)

inventori : CUCITI Vincenzo, GEYMONAT Marina, REMONDINO Chiara  
Lorenza, STABELLINI Barbara, TAMBORRINI Paolo

\*\*\*\*\*

## **DESCRIZIONE**

### **Settore tecnico**

La presente invenzione riguarda in generale il settore delle tecniche di elaborazione dati. In particolare, la presente invenzione riguarda un metodo, un apparato e un programma per elaboratore per elaborare messaggi scambiati su social network.

Nella presente descrizione e nelle rivendicazioni, l'espressione “social network” indicherà un sito web o un programma per elaboratore che consente l'interazione sociale di diversi utenti su internet, in particolare consentendo agli utenti di condividere con i loro contatti personali su Internet informazioni, commenti, messaggi, immagini, ecc.

### **Stato della tecnica**

Oggigiorno, social networks come Twitter® o Facebook® permettono agli utenti di scambiare messaggi su diversi argomenti in scala mondiale. Questi messaggi (ad esempio i “tweets” su Twitter®) possono contenere diverse tipi di contenuti da condividere con altri utenti, ad esempio testi, immagini, audio, video o collegamenti ipertestuali. Ogni messaggio inoltre tipicamente comprende dei metadati associati ai contenuti e comprendenti informazioni

sul messaggio quali l'identificatore del messaggio, username dell'autore del messaggio, data e ora in cui il messaggio è stato pubblicato, luogo dove il messaggio ha avuto origine, ecc.

Sono note tecniche per analizzare messaggi scambiati tramite un social network. Queste analisi sono infatti in grado di fornire informazioni chiave sull'opinione pubblica su diversi argomenti, ad esempio prodotti, opinione politica, intrattenimento, ecc. L'analisi dei messaggi scambiati su social network può inoltre identificare eventi di interesse e fornire informazioni sull'utilità e gli effetti di campagne pubblicitarie nel caso di prodotti, strategie di relazioni pubbliche, figure pubbliche, ecc.

US 9,256,667 descrive un metodo per cercare sorgenti di testo che includono data object ordinati temporalmente, come un blog, che prevede di generare delle query di ricerca basate su termini ed intervalli temporali, identificando data object basandosi sulla query di ricerca e generando delle curve di popolarità basate sulla frequenza di data object che corrispondono a uno o più termini di ricerca in uno o più intervalli di tempo.

### **Riassunto dell'invenzione**

La Richiedente ha percepito che la sopracitata tecnica, descritta in US 9,256,667, ha alcuni inconvenienti.

In particolare la Richiedente ha notato che, secondo quanto descritto in US 9,256,667, quando un certo insieme di messaggi da analizzare è stato identificato applicando una certa query con una o più parole chiave, questo insieme di messaggi identificati viene analizzato da un singolo punto di vista, ossia tenendo conto di una singola caratteristica dei messaggi identificati. US 9,256,667 ad esempio descrive un'analisi temporale, che fornisce un grafico

di popolarità mostrando quante volte una o più parole chiave ricercate sono state menzionate in una particolare finestra temporale; oppure un'analisi della correlazione tra parole chiave, che genera una tabella o un grafico mostrando la correlazione tra le parole chiave ricercate; o un'analisi di "hot keywords", che fornisce una pagina dove ogni parola chiave ricercata è mostrata con una dimensione del carattere proporzionale al suo livello di interesse; o un'analisi geografica, che fornisce una mappa che mostra la distribuzione geografica delle parole chiave ricercate.

Quindi, ogni volta che un utente desidera analizzare da un certo punto di vista i messaggi identificati, deve eseguire delle query separate che producano in uscita un grafico che illustra un singolo aspetto dei messaggi identificati. Se l'utente desidera analizzare da diversi punti di vista i messaggi identificati, svantaggiosamente deve effettuare molte query separate. Questo è scomodo e può richiedere molto tempo.

Ancora peggio, la tecnica descritta in US 9,256,667 non è in grado di fornire un'analisi sistemica dei messaggi identificati, dove con l'espressione "analisi sistemica" si indica un'analisi delle relazioni tra messaggi. La Richiedente ha quindi realizzato che i messaggi scambiati su un social network possono essere reciprocamente legati, ovvero i loro contenuti e/o i loro metadati possono condividere una o più caratteristiche comuni. A titolo di esempio non limitativo, per quanto riguarda il contenuto dei messaggi, le caratteristiche comuni possono essere:

- uno stesso hashtag (ovvero, una parola chiave preceduta dal simbolo "#");
- uno stesso username citato (ovvero, uno username preceduto dal

simbolo "@"

- uno stesso tipo di contenuto (ad esempio, video o audio o immagine);
- una stessa o simile caratteristica nel contenuto (ad esempio immagini con lo stesso soggetto, audio con un pattern tonale simile, ecc.).

Sempre a titolo di esempio non limitativo, per quanto riguarda i metadati, tali caratteristiche comuni possono essere:

- uno stesso autore;
- date e ore di pubblicazione dei messaggi correlati appartenenti allo stesso periodo (ad esempio lo stesso giorno o la stessa settimana);
- luogo di origine dei messaggi correlati appartenente ad una stessa area geografica;
- autori che hanno generato i messaggi correlati che hanno una stessa caratteristica di profilo in comune sul social network (ad esempio numero di contatti, portata, grado di attività, grado di interazione, ecc.)

La Richiedente ha realizzato che un'analisi delle relazioni tra messaggi (e non del singolo messaggio) è più importante ed utile allo scopo di determinare l'andamento e le proprietà delle discussioni su social networks.

Alla luce di quanto sopra, la Richiedente ha risolto il problema di fornire un metodo e un apparato per elaborare i messaggi scambiati su un social network che supera i suddetti problemi.

In particolare, la Richiedente ha affrontato il problema di fornire un metodo e un apparato per elaborare messaggi scambiati su un social network che permetta all'utente di analizzare simultaneamente aspetti multipli delle relazioni tra un certo insieme di messaggi scambiati su un social network.

Secondo forme di realizzazione della presente invenzione, questo

problema è risolto da un apparato per l'elaborazione dati che acquisisce da un social network un insieme di riferimento di messaggi, relativi ad un determinato contesto di riferimento selezionato dall'utente (ad esempio una conferenza, un evento politico o culturale, ecc.). L'apparato applica un'aggregazione multidimensionale all'insieme di riferimento di messaggi, cioè li raggruppa in una serie di array di gruppi avente almeno due dimensioni. In ogni dimensione, l'array comprende un numero di gruppi corrispondente al numero di valori di una certa caratteristica di messaggio (ad esempio in una prima dimensione  $N1$  gruppi corrispondenti a  $N1$  diversi intervalli di tempo  $TS(1)$ ,  $TS(2)$  ...  $TS(N1)$  durante i quali il messaggio è stato pubblicato, in una seconda dimensione  $N2$  gruppi corrispondenti a  $N2$  diversi hashtag  $\#hashtag(1)$ ,  $\#hashtag(2)$ , ...  $\#hashtag(N2)$  compresi nel contenuto del messaggio, e così via). In ogni dimensione, ogni gruppo aggrega tutti i messaggi dell'insieme di riferimento la cui caratteristica di messaggio corrispondente a quella dimensione ha uno stesso valore. Quindi, viene mostrata un'interfaccia utente che è configurata per mostrare l'insieme di riferimento di messaggi come aggregati in almeno due dimensioni dell'array.

Vantaggiosamente, l'aggregazione multidimensionale eseguita dall'apparato di elaborazione dati secondo la presente invenzione permette all'utente di studiare allo stesso tempo le relazioni tra i messaggi presenti nell'insieme di riferimento da vari punti di vista, ovvero di condurre un'analisi sistemica di tali messaggi in modo facile ed intuitivo, interagendo con l'interfaccia utente.

Visto che l'interfaccia utente è configurata in modo da mostrare l'insieme

di riferimento di messaggi come aggregati dall'apparato di elaborazione dati nelle varie dimensioni dell'array, l'utente può ottenere dall'interfaccia utente informazioni su diversi aspetti delle relazioni tra i messaggi dell'insieme di riferimento (ad esempio tempo, hashtag, ecc.) semplicemente interagendo con l'interfaccia utente, senza la necessità di eseguire operazioni complesse come query multiple ecc.

In questo modo, l'utente può identificare andamenti, hashtag chiave, utenti chiave, luoghi, nuovi argomenti, ecc. associati al contesto di riferimento al quale l'insieme di riferimento di messaggi attiene.

Inoltre, l'elaborazione precedente può essere eseguita quasi in tempo reale, ovvero mentre lo scambio di messaggi associati al contesto di riferimento è in corso. L'utente può pertanto eseguire analisi sistemiche di questi messaggi quasi in tempo reale, in modo che lei/lui può utilizzare i risultati dell'analisi per intraprendere azioni all'interno dei contesti di riferimento, ad esempio migliorare l'efficacia delle comunicazioni, aumentare la portata della discussione, aumentare la partecipazione alla discussione di un determinato gruppo di utenti, ecc.

Secondo un primo aspetto, la presente invenzione fornisce un metodo per elaborare messaggi scambiati su un social network da una pluralità di dispositivi utente connessi ad una rete di comunicazione che supporta il social network, il metodo comprendendo, presso un apparato di elaborazione dati cooperante con il social network attraverso la rete di comunicazione:

- a) acquisire dal social network un insieme di riferimento di messaggi riguardanti un contesto di riferimento scelto dall'utente;
- b) applicare una aggregazione multidimensionale all'insieme di riferimento

di messaggi, l'aggregazione multidimensionale comprendendo raggruppare i messaggi dell'insieme di riferimento in un array di gruppi avente almeno due dimensioni, l'array comprendendo, in ogni dimensione, un numero di gruppi che corrispondono ad un numero di valori di una rispettiva caratteristica dei messaggi, in cui ciascun gruppo del numero di gruppi aggrega tutti i messaggi dell'insieme di riferimento la cui rispettiva caratteristica di messaggio ha uno stesso valore, e

- c) mostrare una interfaccia utente configurata per mostrare l'insieme di riferimento di messaggi come aggregati in almeno due dimensioni dell'array.

Preferibilmente, alla fase a) l'insieme di riferimento dei messaggi è ottenuto applicando ai messaggi scambiati sul social network una regola di filtraggio basata su uno o più valori di almeno una caratteristica di messaggio di filtraggio dei messaggi, l'almeno una caratteristica di messaggio di filtraggio essendo scelta tra:

- luogo di origine dei messaggi;
- data e/o intervallo di tempo di pubblicazione dei messaggi;
- autore dei messaggi;
- username e/o hashtag e/o parola chiave citati nei messaggi;
- lingua di messaggi;
- azioni eseguite su messaggi;
- tipo di contenuto di messaggi;
- caratteristiche del contenuto di messaggi.

La regola di filtraggio può essere basata su un singolo valore di una

singola caratteristica di messaggio di filtraggio o su multipli valori di una singola caratteristica di filtraggio. Alternativamente, la regola di filtraggio può essere basata su valori di una combinazione di caratteristiche di messaggio di filtraggio.

Preferibilmente, la fase a) è eseguita da un programma software per l'importazione attraverso l'invocazione di una application programming interface del social network.

Preferibilmente, il metodo inoltre comprende, dopo la fase a), memorizzare l'insieme di riferimento di messaggi in un database accessibile dall'apparato di elaborazione dati, la memorizzazione comprendendo:

- generare un identificatore unico per ogni messaggio dell'insieme di riferimento di messaggi; e/o
- eliminare almeno un campo ridondante o non utilizzato di ciascun messaggio dell'insieme di riferimento di messaggi; e/o
- generare informazioni aggizionali per ogni messaggio dell'insieme di riferimento di messaggi elaborando il contenuto e/o i metadati di ciascun messaggio, le informazioni aggiuntive comprendendo uno o più tra: tipo di contenuto; portata dell'autore del messaggio; e un'indicazione se ciascun messaggio è un messaggio originale o un messaggio ripubblicato.

Preferibilmente, alla fase b) la caratteristica del messaggio è scelta tra:

- luogo di origine dell'insieme di riferimento di messaggi;
- data e/o intervallo di tempo di pubblicazione dell'insieme di riferimento di messaggi;
- autore dell'insieme di riferimento di messaggi;

- username e/o hashtag e/o parole chiave citate nell'insieme di riferimento di messaggi;
- lingua dell'insieme di riferimento di messaggi;
- azioni eseguite sull'insieme di riferimento di messaggi;
- tipo di contenuto dell'insieme di riferimento di messaggi;
- caratteristiche di contenuto dell'insieme di riferimento di messaggi.

Preferibilmente, alla fase b) detto raggruppare comprende creare una raccolta di documenti comprendente una pluralità di oggetti, ogni oggetto corrispondendo ad un rispettivo messaggio dell'insieme di riferimento di messaggi, la raccolta di documenti avente una struttura gerarchica a livelli multipli, ogni livello della struttura gerarchica corrispondendo ad una rispettiva dimensione dell'array.

Preferibilmente, alla fase b) il numero di valori della rispettiva caratteristica del messaggio sono valori predefiniti decisi da un utente.

Preferibilmente, la fase b) comprende inoltre applicare una ulteriore aggregazione multidimensionale all'insieme di riferimento di messaggi, l'ulteriore aggregazione multidimensionale comprendendo l'espansione dell'array di gruppi aggiungendo, in un'ulteriore dimensione, un ulteriore numero di gruppi che corrispondono ad un ulteriore numero di valori, o di cluster di valori, di una caratteristica dei messaggi emergente come rilevata nell'insieme di riferimento di messaggi dall'apparato di elaborazione dati, ogni gruppo dell'ulteriore numero di gruppi aggregando tutti i messaggi dell'insieme di riferimento la cui caratteristica di messaggio emergente ha uno stesso valore dell'ulteriore numero di valori o un valore appartenente ad uno stesso cluster dell'ulteriore cluster di valori.

Preferibilmente, la caratteristica di messaggio emergente è scelta tra:

- autore del messaggio;
- username e/o hashtag e/o parola chiave citata nel messaggio;
- zona di origine del messaggio;
- caratteristica del contenuto del messaggio.

Preferibilmente, un gruppo è aggiunto all'array di gruppi nella dimensione ulteriore se la corrispondente caratteristica è qualificata come significativa, la distribuzione dei valori di questa caratteristica essendo differente da una distribuzione attesa nell'insieme di riferimento di messaggi.

Più preferibilmente, un gruppo è aggiunto all'array di gruppi nella dimensione ulteriore se il numero di occorrenze del valore corrispondente della caratteristica di messaggio emergente nell'insieme di riferimento di messaggi supera una soglia impostata, sotto la quale il corrispondente valore della caratteristica di messaggio emergente è considerata non significativo.

Preferibilmente, la fase c) comprende visualizzare l'interfaccia utente su un display di un terminale connesso con l'apparato di elaborazione dati.

Preferibilmente, l'interfaccia utente è configurata per fornire di default una rappresentazione grafica dell'insieme di riferimento di messaggi come aggregato in almeno una dimensione dell'array, quando l'utente la avvia.

Preferibilmente, l'interfaccia utente è configurata per essere esplorata da detto utente in modo da fornire una rappresentazione grafica di detto insieme di riferimento di messaggi come aggregati in almeno una ulteriore dimensione dell'array.

Secondo un secondo aspetto, la presente invenzione fornisce un apparato di elaborazione dati comprendente un processore configurato per eseguire le

fasi del metodo come riportato sopra.

Secondo un terzo aspetto, la presente invenzione fornisce un programma per elaboratore comprendente istruzioni che, quando eseguite da un elaboratore, inducono l'elaboratore ad eseguire le fasi del metodo come riportato sopra.

#### **Breve descrizione dei disegni**

La presente invenzione risulterà più chiara dalla seguente descrizione dettagliata, fornita a titolo di esempio e non di limitazione, da leggersi con riferimento ai disegni allegati, dove:

- la Figura 1 mostra schematicamente un ambiente esemplificativo dove il metodo per elaborare i messaggi scambiati su un social network è implementato, secondo una forma di realizzazione della presente invenzione;
- la Figura 2 è un diagramma a blocchi delle fasi del metodo per l'elaborazione dei messaggi da un social network, secondo una forma di realizzazione della presente invenzione;
- la Figura 3 mostra schematicamente un'implementazione esemplificativa di alcune fasi del metodo secondo una forma di realizzazione della presente invenzione; e
- le Figure 4a-4e mostrano un'interfaccia utente esemplificativa fornita dal metodo per elaborare messaggi da un social network secondo forme di realizzazione della presente invenzione.

#### **Descrizione dettagliata di forme di realizzazione preferite dell'invenzione**

La Figura 1 mostra schematicamente un ambiente esemplificativo 1 in cui il metodo per l'elaborazione dei messaggi scambiati su un social network è implementato, secondo una forma di realizzazione della presente invenzione.

Detto ambiente esemplificativo 1 comprende una pluralità di dispositivi utente (ad esempio smartphone) 2 connessi ad una rete di comunicazione 3 che supporta il social network 4, in cui "che supporta" significa che il social network 4 sfrutta le connessioni e gli apparati della rete di comunicazione 3 per permettere agli utenti dei dispositivi utente 2 di condividere con i loro contatti personali informazioni, commenti, messaggi, immagini, ecc. Per semplicità, nella Figura 1 solo due dispositivi utente sono mostrati. Tuttavia, è inteso che l'ambiente esemplificativo 1 comprende tipicamente una moltitudine di dispositivi utente 2 (fino a milioni), specialmente se si assume che la rete di comunicazione 3 è internet e il social network 4 è Twitter® o Facebook®.

I dispositivi utente 2 scambiano messaggi 5 (ad esempio tweet, nel caso di Twitter®) attraverso il social network 3, trasmettendo e ricevendo corrispondenti segnali dati verso e dalla rete di comunicazione 3. Le proprietà dei segnali dati trasmessi e ricevuti da ogni dispositivo utente 2 dipendono dal tipo di connessione tra il dispositivo utente 2 e la rete di comunicazione 3, che può essere cablata (ad esempio Ethernet), wireless (ad esempio Wi-Fi) o mobile (ad esempio LTE).

Ogni messaggio 5 scambiato tra i dispositivi utente 2 attraverso il social network 4 preferibilmente comprende un contenuto e dei metadati associati al contenuto.

Per quanto riguarda il contenuto, ogni messaggio 5 può contenere un

testo e/o una immagine e/o un video e/o un audio e/o un collegamento ipertestuale ad un documento esterno (ad esempio una pagina web, un articolo, ecc.). Il testo e/o il documento esterno – se presente – può comprendere ad esempio uno o più hashtag preceduti dal carattere di hash “#” e/o uno o più username citati preceduti dal segno “@”.

Per quanto riguarda i metadati, ogni messaggio 5 tipicamente comprende un identificatore di messaggio unico che identifica il messaggio 5 sul social network 4, uno username dell'autore del messaggio, data e ora in cui il messaggio 5 è stato pubblicato, luogo dove il messaggio 5 ha avuto origine, numero di utenti che hanno già visualizzato il contenuto del messaggio ecc. i messaggi 5 scambiati (ovvero, il loro contenuto e i metadati) sono tipicamente memorizzati in un database del social network 4, non mostrato in Figura 1.

Come mostrato in Figura 1, un apparato di elaborazione dati 6 è preferibilmente connesso al social network 4 tramite la rete di comunicazione 3. L'apparato di elaborazione dati 6 preferibilmente comprende un processore configurato per analizzare i messaggi (e, in particolare, le relazioni tra i messaggi) scambiati dai dispositivi utente 2 sul social network 4, come verrà descritto qui di seguito nel dettaglio con riferimento al diagramma di flusso di Figura 2.

L'apparato di elaborazione dati 6 ha preferibilmente accesso al database 7, che contiene i messaggi da elaborare ed i risultati dell'elaborazione.

Un terminale 8 è inoltre connesso all'apparato di elaborazione dati 6 ed è fornito di un display 80 e mezzi di input 81 (ad esempio tastiera e/o mouse e/o *touchpad* e/o *touchscreen*). Il terminale 8 può essere posizionato in

corrispondenza dell'apparato 6 o può essere un terminale remoto connesso all'apparato 6 attraverso la rete di comunicazione 3. Inoltre, più terminali 8 collegati all'apparato 6 possono essere forniti, in modo da consentire a più utenti di interagire con l'apparato 6.

La Figura 2 è un diagramma di flusso che mostra schematicamente le fasi del metodo per l'elaborazione dei messaggi 5 scambiati sul social network 4, secondo una forma di realizzazione della presente invenzione, come eseguita dal processore dell'apparato di elaborazione dati 6.

Una prima fase 101 comprende acquisire dal social network 4 un insieme di riferimento di messaggi 5 selezionati tra tutti i messaggi 5 scambiati sul social network 4. Alla fase 101 l'apparato di elaborazione 6 può ad esempio applicare un filtro che l'utente configura per selezionare, tra tutti i messaggi 5 scambiati sul social network 4, solo quelli riguardanti un certo contesto di riferimento, come ad esempio uno sport o evento culturale, iniziativa aziendale, conferenza, ecc.

Come regola di filtraggio per configurare il filtro da applicare alla fase 101, l'utente può utilizzare una o più caratteristiche del contenuto di messaggio e/o i metadati associati ad esso, ad esempio scegliendo tra:

- a) luogo di origine dei messaggi;
- b) data e/o intervallo di tempo di pubblicazione dei messaggi;
- c) autore/i dei messaggi;
- d) username citati nel contenuto dei messaggi;
- e) hashtag contenuti nel contenuto dei messaggi;
- f) parole chiave (ad esempio nome di una città, di una celebrità o personaggio, di un evento, ecc.) contenuti nel contenuto dei messaggi;

- g) lingua dei messaggi;
- h) azioni eseguite sui messaggi all'interno del social network 4 (pubblicazioni, condivisioni, like, inoltro, retweeting, citazioni, ecc.);
- i) tipo di contenuto dei messaggi (testo, immagine, video, collegamento ipertestuale ad un contenuto esterno, ecc.);
- j) caratteristica del contenuto (ad esempio aspetto o soggetto di un'immagine, pattern tonale di una traccia audio, ecc.).

La regola di filtraggio può comprendere un singolo valore di una caratteristica (ad esempio un singolo hashtag, un singolo autore del messaggio, un singolo tipo di contenuto, ecc.) oppure valori multipli di una stessa caratteristica (ad esempio hashtag multipli collegati, multipli autori, ecc.). In ogni caso, in questa fase il/i valore/i della/delle caratteristica/caratteristiche da utilizzare dall'apparato di elaborazione dati 6 per filtrare i messaggi 5 e identificare l'insieme di riferimento di messaggi 50 sono decisi dall'utente a priori, cioè senza alcuna conoscenza del contenuto dei messaggi o dei metadati.

Come regola per il filtraggio, può essere usata una singola caratteristica, o una loro combinazione. Ad esempio, se una analisi deve essere condotta sui messaggi 5 scambiati sul social network 4 in relazione ad un evento od un argomento identificato da uno o più hashtag collegati, la regola di filtraggio può essere solo su tale/i hashtag, in modo che alla fase 101 il filtro applicato dall'apparato di elaborazione dati 6 selezioni unicamente i messaggi comprendenti tale/i hashtag. Se una analisi deve essere eseguita sui messaggi 5 scambiati sul social network 4 ad esempio per scoprire le attività e gli eventi nel fine settimana in una certa città e indagare la loro popolarità,

la regola di filtraggio può essere una combinazione di data ed intervallo di tempo di pubblicazione dei messaggi e luogo d'origine dei messaggi, in modo che alla fase 101 il filtro applicato dall'apparato di elaborazione dati 6 selezioni solo i messaggi 5 pubblicati durante il fine settimana e generati nella città di interesse. Se si desiderasse limitare ulteriormente quest'ultima analisi ad esempio alle sole fotografie, la regola di filtraggio precedente può essere ulteriormente combinata con il tipo di contenuto dei messaggi, in modo che alla fase 101 il filtro selezioni solo i messaggi contenenti fotografie. E così via.

Per esempio, se il social network 4 è Twitter®, la fase 101 può essere eseguita dall'apparato di elaborazione dati 6 usando un noto programma software di importazione acquisendo i messaggi utilizzando una invocazione dell'API di streaming di Twitter® (Public Stream, endpoint POST statuses/filter", parameter "track") che produce una notifica in formato JSON (JavaScriptObject Notation) ogni volta che un messaggio 5 che rispetta la regola di filtraggio è trovato. Opzionalmente, l'importatore può invocare uno o più moduli software ad-hoc configurati per arricchire i metadati dei messaggi scambiati sul social network 4 con informazioni che non sono disponibili per l'API e che sono parte delle regole di filtraggio applicabili alla fase 101 (ad esempio la caratteristica del contenuto al precedente punto j)).

Una volta che l'insieme di riferimento di messaggi 50 è stato acquisito dall'apparato di elaborazione dati 6 dal social network 4 applicando i filtri descritti in precedenza, l'apparato di elaborazione dati 6 preferibilmente salva ogni messaggio dell'insieme di riferimento 50 nel database 7 (fase 102).

A tale scopo, alla fase 102 l'apparato di elaborazione dati 6

preferibilmente prepara ciascun messaggio dell'insieme di riferimento 50 per il salvataggio. In particolare, alla fase 102 un identificatore unico è preferibilmente generato per ogni messaggio da salvare, che identifica univocamente il messaggio nel contesto dell'elaborazione eseguita dall'apparato di elaborazione dati 6 ( e che in generale è diverso da – e più corto di - quello che identifica il messaggio all'interno del social network 4). L'identificatore di messaggio unico generato alla fase 102 può essere ad esempio una stringa alfanumerica.

Inoltre, alla fase 102, il formato di ogni messaggio da salvare è preferibilmente cambiato in un formato adatto al salvataggio. Ad esempio, i formati dei timestamp sono adattati, i campi ridondanti od inutilizzati di ciascun messaggio vengono cancellati, ecc.

Inoltre, alla fase 102, per ogni messaggio da salvare vengono preferibilmente generate informazioni aggiuntive, elaborando il contenuto del messaggio e/o i metadati del messaggio. Ad esempio, tali informazioni aggiuntive possono comprendere:

- il tipo di contenuto (video, immagine, audio, collegamento ipertestuale ad un documento) compreso nel messaggio;
- la portata (ovvero, il numero di contatti sul social network 4) dell'autore del messaggio; e
- un'indicazione se il messaggio è originale oppure se è un messaggio ripubblicato (ad esempio un tweet o un retweet, nel caso in cui il social network 4 sia Twitter®).

Dopo tale elaborazione, alla fase 102 ogni messaggio dell'insieme di riferimento 50 è quindi salvato nel database 7, assieme al suo identificatore

univoco, al suo contenuto, ai suoi metadati e alle sue informazioni aggiuntive.

Dopo di che, l'apparato di elaborazione dati 6 preferibilmente applica una prima aggregazione multidimensionale ai messaggi dell'insieme di riferimento 50 (fase 103).

La prima aggregazione multidimensionale preferibilmente comprende raggruppare i messaggi dell'insieme di riferimento 50 in un array multidimensionale di gruppi 9. In ogni dimensione, l'apparato di elaborazione dati 6 fornisce l'array 9 con un numero di gruppi che corrispondono ad un numero di valori predefiniti (ovvero, i valori scelti dall'utente) di una certa caratteristica di messaggio, in particolare del contenuto dei messaggi, dei metadati o delle informazioni aggiuntive. In particolare, in una dimensione corrispondente ad una certa caratteristica di messaggio, ogni gruppo dell'array 9 aggrega tutti i messaggi dell'insieme di riferimento 50 la cui caratteristica ha uno stesso valore tra un numero di valori predefiniti stabiliti dall'utente.

Le caratteristiche dei messaggi considerate in detto primo raggruppamento multidimensionale alla fase 103 possono essere scelte tra le caratteristiche a) - j) elencate in precedenza.

Ad esempio, assumendo che l'apparato di elaborazione dati 6 abbia selezionato un insieme di riferimento di messaggi 50 basandosi su un hashtag primario `#primary_hashtag` associato ad un certo evento, l'apparato di elaborazione dati 6 può quindi raggruppare i messaggi dell'insieme di riferimento 50 in un array multidimensionale 9 schematicamente raffigurato in Figura 3 e comprendente:

- N1 gruppi in una prima dimensione, ogni gruppo comprendendo tutti i messaggi dell'insieme di riferimento 50 che sono stati pubblicati nello stesso intervallo di tempo di N1 intervalli di tempo  $TS(1)$ ,  $TS(2)$  ...  $TS(N1)$  decise dall'utente (preferibilmente, l'utente può selezionare il momento di inizio e la durata di ogni intervallo di tempo);
- N2 gruppi in una seconda dimensione, ogni gruppo comprendendo tutti i messaggi dell'insieme di riferimento 50 contenenti lo stesso hashtag secondario di N2 diversi hashtag secondari  $\#hashtag(1)$ ,  $\#hashtag(2)$ , ...  $\#hashtag(N2)$  collegati all'hashtag primario  $\#primary\_hashtag$  e decisi dall'utente.
- N3 gruppi in una terza dimensione, ogni gruppo comprendendo tutti i messaggi dell'insieme di riferimento 50 (ovvero, comprendenti l'hashtag  $\#primary\_hashtag$ ) contenenti lo stesso tipo di contenuto di N3 tipi di contenuto  $type(1)$ ,  $type(2)$ , .. ,  $type(N3)$  decisi dall'utente; e così via.

Al fine di formare ogni singolo gruppo dell'array 9 basandosi su un certo valore di una certa caratteristica di messaggio, l'apparato di elaborazione dati 6 può usare qualsiasi tecnica nota. Se la caratteristica si trova nei metadati associati al contenuto dei messaggi, ad esempio, è possibile utilizzare le tecniche descritte da US 2013/0036117 o US 2006/0036462. Se la caratteristica è una caratteristica testuale nel contenuto dei messaggi, la tecnica descritta da US 9,256,667 può essere utilizzata.

Si noti che un gruppo in una certa dimensione dell'array 9 è vuoto, se nessuno dei messaggi dell'insieme di riferimento 50 ha la caratteristica di messaggio corrispondente a quella dimensione con il predefinito valore

corrispondente a quel gruppo. Inoltre, l'apparato di elaborazione dati 6 può inserire uno stesso messaggio dell'insieme di riferimento 50 simultaneamente in più gruppi. Ad esempio, un messaggio comprendente entrambi gli hashtag secondari #hashtag(1), #hashtag(2) e che è stato pubblicato durante l'intervallo temporale TS(N1) è compreso simultaneamente in tutti i gruppi relativi a tali valori delle caratteristiche.

Per implementare la sopra descritta aggregazione multidimensionale sull'insieme di riferimento di messaggi 50, l'apparato di elaborazione dati 6 preferibilmente etichetta opportunamente i messaggi dell'insieme di riferimento 50. In particolare, l'apparato di elaborazione dati 6 crea una raccolta di documenti, avente una struttura gerarchica con più livelli, ogni livello corrispondendo ad una dimensione dell'array 9.

Ad esempio, considerando il precedente array 9 esemplificativo, una porzione di una esemplificativa collezione di documenti scritti in JSON è riportata qui di seguito:

```
{
  "#hashtag(1)": {
    "TS(1)": [obj1, obj2, obj3],
    "TS(2)": [obj4, obj5, obj6],
    ....
    "TS(N1)": [...],
  },
  "#hashtag(2)": {
    "TS(1)": [...],
    "TS(2)": [...],
    ....
    "TS(N1)": [...],
  },
  ...
}
```

```
"#hashtag(N2)": {  
    "TS(1)": [...],  
    "TS(2)": [...],  
    ....  
    "TS(N1)": [...],  
    },  
}
```

Ogni singolo oggetto indicato tra parentesi è essenzialmente un puntatore all'identificatore del messaggio ed una indicazione del tipo di contenuto del messaggio, vale a dire:

```
obj: {  
    "id": "id1",  
    "type": "image"  
}
```

Quindi, secondo forme di realizzazione preferite, l'apparato di elaborazione dati 6 applica una seconda aggregazione multidimensionale all'insieme di riferimento di messaggi 50 (fase 104). La seconda aggregazione multidimensionale preferibilmente comprende, da parte dell'apparato di elaborazione dati 6, espandere l'array multidimensionale 9 creato alla fase 103 aggiungendo nuovi gruppi in almeno una dimensione ulteriore, e raggruppare nuovamente i messaggi dell'insieme di riferimento 50 in tali nuovi gruppi.

In particolare, preferibilmente, alla fase 104 l'apparato di elaborazione dati 6 preferibilmente legge il contenuto e/o i metadati di ogni messaggio dell'insieme di riferimento 50 per scoprire una o più caratteristiche di messaggio significative ulteriori. Tali caratteristiche di messaggio ulteriori (definite da qui in avanti "caratteristiche emergenti") possono essere, ad esempio:

- k) ulteriori autori/e dei messaggi, diversi da quelli che potrebbero essere stati usati alla fase 101 o 103;
- l) ulteriori username citati nel contenuto dei messaggi, diversi da quelli che potrebbero essere stati usati alla fase 101 o 103;
- m) ulteriori hashtag contenuti nel contenuto dei messaggi, diversi da quelli che potrebbero essere stati usati alla fasi 101 o 103;
- n) ulteriori parole chiave (ad esempio nome di una città, di una celebrità o di una personalità, di un evento, ecc.) contenute nel contenuto dei messaggi, diverse da quelle che potrebbero essere state usate alla fase 101 o 103.
- o) ulteriore caratteristica di contenuto (ad esempio soggetto di un'immagine, pattern tonale di un audio, ecc.) diversa da quelle che potrebbero essere state utilizzate alle fase 101 o 103.

Preferibilmente, per eseguire la fase 104 il sistema è impostato preselezionando l'insieme di caratteristiche tra le quali l'apparato di elaborazione dati 6 cerca le caratteristiche emergenti da scoprire nei messaggi dell'insieme di riferimento 50 e definendo le distribuzioni attese di detto insieme di caratteristiche. Preferibilmente l'utente può impostare un numero massimo di caratteristiche emergenti che possono essere visualizzate, ma non ha una conoscenza a priori di quali saranno le caratteristiche emergenti e i valori di tali caratteristiche emergenti nei messaggi. L'utente inoltre non sa neanche se quella caratteristica emergente è presente o meno nei messaggi dell'insieme di riferimento 50. Più preferibilmente, tuttavia, il numero massimo di caratteristiche emergenti da visualizzare è impostato dall'utente a 2.

Inoltre, l'utente può decidere di visualizzare caratteristiche specifiche dell'insieme di riferimento 50 anche se non si qualificano come caratteristiche emergenti in base alla descrizione precedente. Ad esempio, l'utente può visualizzare ulteriori hashtag contenuti nel contenuto dei messaggi, ma non sa a priori quali tali ulteriori hashtag possano essere o addirittura se hashtag diversi da quelli utilizzati dall'apparato di elaborazione dati 6 per la fase di filtraggio 101 o per la prima fase di raggruppamento 103 sono presenti nei messaggi.

Per ogni caratteristica emergente da scoprire, alla fase 104 l'apparato di elaborazione dati 6 aggiunge ulteriori gruppi all'array 9 creato alla fase 103, in una sua corrispondente ulteriore dimensione .

Per fare in modo che una caratteristica possa qualificarsi come significativa, e quindi come una "caratteristica emergente", la distribuzione dei valori di questa caratteristica nell'insieme di riferimento di messaggi 50 deve essere diversa da una distribuzione attesa. La distribuzione attesa dipende dalla specifica caratteristica: in alcuni casi è una distribuzione casuale, in altri casi è ben nota, ad esempio la distribuzione giornaliera del traffico dati è una curva che aumenta al mattino e diminuisce di notte. In caso di una caratteristica che è rappresentata tramite un insieme di parametri (tipicamente chiamate caratteristiche di basso livello, ovvero caratteristiche con un valore semantico basso, come una immagine o un video rappresentato attraverso luminanza, istogrammi dei colori o texture) la deviazione dalla distribuzione attesa può essere identificata eseguendo un raggruppamento multidimensionale o clustering, in modo da identificare la presenza di gruppi o cluster particolarmente densi o rilevanti, che è segno

della deviazione dalla distribuzione prevista. Maggiore è la deviazione dalla distribuzione attesa, maggiore è la rilevanza della caratteristica corrispondente. La deviazione dalla distribuzione attesa, ovvero, può contenere informazioni significative.

Quindi, per ogni caratteristica emergente, per ogni diverso valore, o cluster di valori, identificati come significativamente scostantisi dalla distribuzione attesa, l'apparato di elaborazione dati 6 crea un gruppo separato, che raggruppa tutti i messaggi dell'insieme di riferimento 50 che comprendono la caratteristica emergente con quel valore o con un valore appartenente a quel cluster.

Ad esempio, sempre in riferimento alla Figura 3, se l'array multidimensionale alla fine della fase 103 ha N dimensioni, alla fase 104 l'apparato di elaborazione dati 6 può aggiungere i seguenti gruppi:

- N1' gruppi nella (N+1)esima dimensione, ogni gruppo comprendendo tutti i messaggi dell'insieme di riferimento 50 aventi lo stesso autore `author(1)`, `author(2)`, ... `author(N1')` come scoperto nei metadati dei messaggi dell'insieme di riferimento 50;
- N2' gruppi in una (N+2)esima dimensione, ogni gruppo comprendendo tutti i messaggi dell'insieme di riferimento 50 contenenti lo stesso username citato `@username(1)`, `@username(2)`, ... `@username(N2')` come scoperto nel contenuto dei messaggi dell'insieme di riferimento 50;
- N3' gruppi in una (N+3)esima dimensione, ogni gruppo comprendendo tutti i messaggi dell'insieme di riferimento 50 contenenti uno stesso hashtag emergente `#ehashtag(1)`, `#ehashtag(2)`, ...

#ehashtag(N3') come scoperto nel contenuto dei messaggi dell'insieme di riferimento 50;

- N4' gruppi in una (N+4)esima dimensione, ogni gruppo comprendendo tutti i messaggi dell'insieme di riferimento 50 originati da azioni di diversi tipi (pubblicazione, inoltro, condivisione, ripubblicazione, ecc.) su uno stesso messaggio originale, come scoperto nelle informazioni aggiuntive dei messaggi dell'insieme di riferimento 50 (tali gruppi non sono mostrati in Figura 3 per semplicità); e così via.

Secondo una forma di realizzazione preferita, l'apparato di elaborazione dati 6 crea un gruppo per un certo valore di una caratteristica emergente se il numero di occorrenze di quel valore nei messaggi dell'insieme di riferimento 50 supera una soglia preimpostata, sotto la quale quel valore della caratteristica emergente è considerato non significativo. L'utente può impostare la soglia per ogni caratteristica emergente da scoprire tenendo conto, ad esempio, del tipo di caratteristica emergente e del numero totale di messaggi nell'insieme di riferimento 50. Ad esempio, una singola occorrenza di un certo autore può essere considerata sufficiente per creare un gruppo corrispondente, mentre sono preferibili diverse occorrenze di uno stesso hashtag emergente o username citato per considerare l'hashtag emergente o lo username citato statisticamente significativi e creare un gruppo corrispondente nell'array 9.

Le aggregazioni alle fasi 103 e 104 possono essere eseguite da un processo autonomo (ad esempio un processo Node.js), possibilmente usando ulteriori dati accessibili attraverso una interfaccia di programmazione (ad esempio Web API).

L'apparato di elaborazione dati 6 quindi salva preferibilmente l'array di gruppi 9 nel database 7 (fase 105).

Una volta che l'apparato di elaborazione dati 6 ha completato l'array di gruppi 9 come descritto precedentemente e lo ha memorizzato sul database 7, in qualsiasi momento l'utente può utilizzare il terminale 8 collegato con l'apparato di elaborazione dati 6 per visualizzare una interfaccia grafica 10 che mostra i risultati elaborati (fase 106). In particolare, l'interfaccia utente 10 è preferibilmente visualizzata sul display 80 del terminale 8.

Come verrà descritto in dettaglio qui di seguito con riferimento alle Figure 4a-4e, l'interfaccia utente 10 offre vantaggiosamente una vista sistemica dell'insieme di riferimento di messaggi 50, ovvero una vista delle relazioni tra questi messaggi. L'interfaccia utente 10 in particolare è configurata per fornire una rappresentazione grafica dei messaggi come aggregati in ogni dimensione dell'array 9, ogni dimensione corrispondendo ad un certo aspetto delle relazioni tra i messaggi.

Una rappresentazione grafica dei messaggi come aggregati in almeno una dimensione dell'array 9 è preferibilmente fornita di default dall'interfaccia utente 10, non appena l'utente la apre sul display 80 del terminale 8. Tramite i mezzi di input 81 del terminale 8, l'utente può richiedere all'interfaccia utente 10 di fornire una rappresentazione grafica dei messaggi come aggregati in altre dimensioni dell'array 9, a seconda dell'aspetto/i delle relazioni tra i messaggi che desidera indagare.

una interfaccia utente 10 esemplificativa è illustrata schematicamente nelle Figure 4a-4e. La Figura 4a mostra l'aspetto iniziale dell'interfaccia

utente 10, ovvero il suo aspetto quando l'utente l'avvia, prima che inizi ad interagire con essa.

Come mostrato in Figura 4a, l'interfaccia utente 10 può comprendere informazioni indicative dell'insieme di riferimento di messaggi 50 nel suo insieme, ad esempio un titolo 20 indicativo del contesto di riferimento considerato ed un indicatore 21 del numero totale di messaggi nell'insieme di riferimento 50.

L'interfaccia utente 10 inoltre fornisce preferibilmente un indicatore dell'insieme di riferimento di messaggi 50 come aggregato in una prima dimensione dell'array 9, in particolare la dimensione corrispondente al tipo di contenuto  $type(1)$ ,  $type(2)$ , ...  $type(N3)$ . Assumendo che solo tre tipi di contenuti siano considerati (ad esempio documenti, immagini e video), per ogni tipo di contenuto l'interfaccia utente 10 preferibilmente visualizza un indicatore 22 del numero di messaggi comprendenti quel tipo di contenuto. Ogni tipo è preferibilmente associato ad una rispettiva icona.

L'interfaccia utente 10 inoltre comprende un grafico 23 che mostra l'insieme di riferimento di messaggi 50 come aggregato in una seconda dimensione dell'array 9, in particolare la dimensione corrispondente agli intervalli di tempo  $TS(1)$ ,  $TS(2)$  ...  $TS(N1)$  nei quali i messaggi sono stati pubblicati. A titolo di esempio non limitativo, il grafico 23 comprende una linea temporale (in particolare, una linea temporale curva) dalla quale un numero di raggi 24 è proiettato verso l'esterno. Ogni raggio 24 corrisponde ad un corrispettivo intervallo di tempo  $TS(1)$ ,  $TS(2)$  ...  $TS(N1)$  e ha una lunghezza indicativa del numero di messaggi pubblicati in quell'intervallo di tempo. In particolare, ogni raggio 24 è formato da un numero di elementi

uguale al numero di messaggi pubblicati durante il corrispondente intervallo di tempo.

Il grafico 23 è inoltre configurato per mostrare l'insieme di riferimento di messaggi 50 come aggregati in una seconda dimensione dell'array 9, in particolare la dimensione corrispondente agli hashtag secondari  $\#hashtag(1)$ ,  $\#hashtag(2)$ , ...  $\#hashtag(N2)$ . Ad esempio, ogni hashtag secondario può essere associato ad un rispettivo colore come indicato da una legenda 25 anch'essa mostrata dall'interfaccia utente 10. All'interno di ciascun raggio 24, ogni elemento è quindi colorato col colore associato all'hashtag secondario che contiene.

All'interno di ciascun raggio 24, ciascun elemento rappresentante un messaggio preferibilmente ha proprietà grafiche che forniscono una visione grafica e immediata di alcune caratteristiche del messaggio stesso. Ad esempio, due diversi tipi di elementi (ad esempio cerchi e stelle, come mostrato in Figura 4a) o due diversi colori possono essere usati per rappresentare rispettivamente i messaggi originali (ad esempio i tweet) e i messaggi ripubblicati (ad esempio i retweet), come indicato da una ulteriore legenda 26 anch'essa compresa nell'interfaccia utente 10.

Inoltre, ogni elemento rappresentante un messaggio può essere cerchiato da un rispettivo cerchio 27 (solo due cerchi 27 in un singolo raggio 24 sono mostrati in Figura 4a, per semplicità) che fornisce una indicazione della portata del messaggio all'interno del social network 4. Diversi diametri dei cerchi 27 indicano diversi intervalli del valore della portata, come indicato da una ulteriore legenda 28 anch'essa compresa nell'interfaccia utente 10. L'interfaccia utente 10 può inoltre comprendere, per ogni raggio 24, un

rispettivo ulteriore cerchio 29 che indica la portata complessiva, sul social network 4, di tutti i messaggi pubblicati durante l'intervallo di tempo associato. La portata complessiva rappresentata dal cerchio 29 è calcolata unendo le portate di tutti i messaggi dell'intervallo di tempo.

Pertanto, anche prima che l'utente inizi ad interagire con l'interfaccia utente 10, come impostazione predefinita l'interfaccia utente 10 offre vantaggiosamente una rappresentazione molto intuitiva di diversi aspetti delle relazioni tra i messaggi dell'insieme di riferimento 50 (tempo, hashtag secondari, tipo di contenuto), ogni aspetto corrispondendo ad una diversa dimensione dell'array di gruppi 9 risultante dalle aggregazioni eseguite dell'apparato di elaborazione dati 6 alle fasi 103-104. L'interfaccia utente 10 inoltre offre vantaggiosamente di default una rappresentazione davvero intuitiva di alcune caratteristiche di ogni singolo messaggio (tweet o retweet, portata).

Inoltre, come detto precedentemente, l'utente può interagire con l'interfaccia utente 10 in modo da investigare altri aspetti delle relazioni tra i messaggi, che sono collegati ad altre dimensioni dell'array multidimensionale 9.

Ad esempio, se l'utente seleziona una delle icone corrispondenti ad un tipo di contenuto  $type(1)$ ,  $type(2)$ , ...  $type(N3)$ , l'interfaccia utente 10 preferibilmente mostra in ulteriore dettaglio l'insieme di riferimento di messaggi 50 come aggregati nella dimensione dell'array 9 che corrisponde al tipo di contenuto. In particolare, come mostrato in Figura 4b, se l'utente seleziona una certo tipo di contenuto (ad esempio video) per mezzo di un puntatore P, l'interfaccia utente 10 mostra tutti i messaggi (tweet e retweet)

contenenti video. Ad esempio, tutti gli elementi del raggio del grafico 23 corrispondenti a messaggi contenenti tipi di contenuti diversi da quello selezionato possono essere sbiaditi, in modo che solo questi ultimi sono evidenziati, come mostrato in Figura 4b. Questo fornisce una vista immediata ed intuitiva dei messaggi contenenti video.

Preferibilmente, anche l'indicatore 21 è aggiornato, in modo da indicare il numero complessivo di messaggi (tweet e retweet) contenenti unicamente il tipo di contenuto selezionato.

In aggiunta, l'interfaccia utente 10 può inoltre mostrare l'insieme di riferimento di messaggi 50 come aggregati in un'ulteriore dimensione dell'array 9, in particolare quella dei messaggi originati da azioni di tipi differenti (pubblicazioni, inoltro, condivisione, ripubblicazione, ecc.) su uno stesso messaggio. Se, partendo dall'interfaccia utente 10 come in Figura 4a, l'utente seleziona un messaggio (ad esempio un tweet) cliccando il corrispondente elemento del raggio nel grafico 23, l'interfaccia utente 10 mostra tutti i messaggi dell'insieme di riferimento 50 che sono collegati al messaggio selezionato da azioni di diversi tipi (ad esempio i suoi retweet). Ad esempio, come mostrato in Figura 4c, tutti gli elementi del raggio del grafico 23 – ad eccezione del messaggio selezionato e dei messaggi ad esso collegati, assieme ai loro rispettivi cerchi di portata – possono essere sbiaditi, in modo che solo questi ultimi siano evidenziati. Questo fornisce un'immediata ed intuitiva visione dell'effetto "passaparola" del messaggio selezionato. Ad esempio, dalla situazione esemplificativa illustrata nella Figura 4c, l'utente può vedere che un tweet selezionato da un puntatore P – che era stato pubblicato durante il primo intervallo di tempo considerato – è

stato retweetato un certo numero di volte all'interno dello stesso intervallo di tempo e successivamente un numero decrescente di volte nei successivi intervalli di tempo. Nessun retweet è stato fatto per un certo numero di intervalli di tempo (probabilmente perché l'argomento della discussione era cambiato), ma successivamente una raffica di retweet viene nuovamente trovata durante un certo periodo di tempo (probabilmente perché la discussione è tornata su un argomento collegato a quello del tweet originale).

Preferibilmente, anche gli indicatori 21 e 22 vengono aggiornati, in modo che indichino il numero complessivo di messaggi collegati a quello selezionato e, per ogni tipo di contenuto, il numero di messaggi contenenti quel tipo di contenuto.

Allo stesso tempo, come schematicamente mostrato in Figura 4c, quando l'utente seleziona uno specifico messaggio (sia un tweet o un retweet) cliccando il corrispondente elemento del raggio nel grafico 23, viene aperta una finestra 30 che mostra informazioni sul messaggio, che includono:

- author 300 del messaggio; e
- anteprima 301 del contenuto del messaggio.

Selezionando l'anteprima 301, l'utente può accedere al contenuto completo del messaggio selezionato.

Selezionando author 300, invece, l'interfaccia utente 10 preferibilmente mostra l'insieme di riferimento di messaggi 50 come aggregato in un'ulteriore dimensione dell'array 9, ossia la dimensione corrispondente all'autore del messaggio. In particolare, come illustrato in Figura 4d, l'interfaccia utente 10 mostra tutti i messaggi dello stesso autore del messaggio selezionato. Ad esempio, tutti gli elementi del raggio del grafico 23, ad eccezione di quelli di

author, possono essere sbiaditi, in modo che solo questi ultimi siano evidenziati, come mostrato in Figura 4d. Questo fornisce una immediata ed intuitiva visione della partecipazione di author alla discussione. Ad esempio, dalla situazione esemplificativa mostrata in Figura 4d, l'utente può vedere che author è stato attivo per un certo numero di intervalli di tempo (probabilmente perché l'argomento della discussione era di suo interesse), successivamente è rimasto in silenzio per un certo numero di intervalli di tempo e successivamente ha ricominciato a pubblicare (probabilmente perché la discussione è tornata su un argomento di suo interesse).

Preferibilmente, anche gli indicatori 21 e 22 sono aggiornati, in modo che indichino il numero complessivo di messaggi pubblicati o ripubblicati da author e, per ogni tipo di contenuto, il numero di messaggi contenenti quel tipo di contenuto.

L'utente può quindi esplorare ulteriormente uno qualsiasi dei messaggi evidenziati in Figura 4d, selezionando il suo elemento di raggio associato 24. Quando l'utente seleziona uno qualsiasi di tali messaggi, una finestra simile alla finestra 30 è aperta mostrando informazioni sul messaggio selezionato. In questo modo, l'utente può navigare facilmente tra i messaggi di author ed analizzarli con il desiderato livello di dettaglio. Ad esempio, analizzando il contenuto dei messaggi di author, l'utente può capire le ragioni dei cambiamenti del suo livello di partecipazione alla discussione.

L'interfaccia utente 10 inoltre comprende preferibilmente una lista di hashtag emergenti. 31 ed una lista di citazioni 32

La lista di hashtag emergenti 31 comprende preferibilmente gli hashtag emergenti #ehashtag(1), #ehashtag(2), ... #ehashtag(N3')

scoperti nel contenuto dei messaggi dell'insieme di riferimento 50 durante la seconda aggregazione alla fase 104 e per i quali un rispettivo gruppo è stato creato nell'array multidimensionale 9. Gli hashtag emergenti  $\#e\text{hashtag}(1)$ ,  $\#e\text{hashtag}(2)$ , ...  $\#e\text{hashtag}(N3')$  come mostrati dall'interfaccia utente 10 sono preferibilmente ordinati in base al numero di occorrenze nell'insieme di riferimento di messaggi 50. A titolo di esempio, nelle figure 4a-4e si assume che  $\#e\text{hashtag}(2)$  sia l'hashtag emergente più popolare.

Allo stesso modo, la lista di citazioni 32 comprende preferibilmente gli username citati  $@\text{username}(1)$ ,  $@\text{username}(2)$ , ...  $@\text{username}(N2')$  come scoperti nel contenuto dei messaggi dell'insieme di riferimento 50 durante la seconda aggregazione alla fase 104 e per i quali un rispettivo gruppo è stato creato nell'array multidimensionale 9.

Anche gli username citati  $@\text{username}(1)$ ,  $@\text{username}(2)$ , ...  $@\text{username}(N2')$  come mostrati dall'interfaccia utente 10 sono preferibilmente ordinati in base al numero di occorrenze nell'insieme di riferimento di messaggi 50. A titolo di esempio, nelle figure 4a-4e si assume che  $@\text{username}(N2')$  sia la citazione più popolare.

Le liste 31 e 32 forniscono all'utente un'immediata visione degli hashtag emergenti e delle citazioni come scoperti alla fase 104 nel contenuto dei messaggi dell'insieme di riferimento 50. Questo permette all'utente di scoprire hashtag e username (ad esempio quelli di celebrità o figure pubbliche) che i partecipanti alla discussione hanno ritenuto essere rilevanti all'interno del contesto di riferimento considerato.

Oltretutto, utilizzando la lista 31 o 32 l'utente può utilizzare l'interfaccia utente 10 per mostrare l'insieme di riferimento di messaggi 50 come aggregato nelle ulteriori dimensioni dell'array 9, in particolare quelle degli hashtag emergenti o delle citazioni o delle zone di origine dei messaggi. In particolare, selezionando un hashtag emergente dalla lista 31 o una citazione dalla lista 32, l'interfaccia utente 10 preferibilmente evidenzia nel grafico 23 i messaggi che contengono quel hashtag emergente o citazione. A titolo di esempio non limitativo, come mostrato in Figura 4e, se l'utente seleziona #ehtag(2) dalla lista 31, un numero di frecce 33 è mostrato dall'interfaccia utente 10, ogni freccia connettendo #ehtag(2) selezionato con l'elemento del raggio che rappresenta un messaggio contenente #ehtag(2).

Preferibilmente, come mostrato in Figura 4e, anche gli indicatori 21 e 22 vengono aggiornati, in modo che indichino il numero complessivo di messaggi contenenti il selezionato hashtag emergente o citazione e, per ogni tipo di contenuto, il numero di messaggi contenente quel tipo di contenuto.

L'interfaccia utente 10 può inoltre comprendere un campo di ricerca di testo 34 per cercare messaggi contenenti una certa parola selezionata dall'utente.

Preferibilmente, l'utente ha la possibilità di personalizzare l'interfaccia utente 10. A questo scopo, prima dell'apertura dell'interfaccia utente 10 (per la prima volta o per ogni volta successiva, se desidera effettuare delle modifiche), l'utente preferibilmente compila e un file di configurazione dove può indicare:

- inizio e fine dell'intervallo di tempo complessivo da visualizzare nel grafico 23 dell'interfaccia utente 10.
- geolocalizzazione dei messaggi da mostrare nell'interfaccia utente 10
- hashtag secondari che, tra # hashtag(1), #hashtag(2), ... #hashtag(N2), sono di particolare interesse e devono quindi essere mostrati (ad esempio attraverso l'uso di colori differenti degli elementi del raggio 24, come descritto precedentemente) nel grafico 23 dell'interfaccia utente 10. Il numero di hashtag secondari che può essere visualizzato è preferibilmente minore di una soglia predefinita (ad esempio 12);
- inizio e fine di ogni intervallo di tempo del grafico 23 (ogni raggio del grafico 23 può aggregare messaggi da intervalli di tempo multipli consecutivi TS(1), TS(2) ... TS(N1));
- tipo di messaggi (ad esempio solo tweet o entrambi tweet e retweet) da visualizzare nel grafico 23;
- se desidera che gli hashtag emergenti 31 e/o le citazioni 32 siano mostrati dall'interfaccia utente 10 e, in caso affermativo, il numero di hashtag emergenti e/o citazioni da mostrare, che è preferibilmente minore di una soglia massima predefinita (ad esempio 30); e
- i colori da associare agli hashtag secondari ed allo sfondo dell'interfaccia utente 10.

Inoltre, vantaggiosamente, l'aggregazione multidimensionale eseguita dall'apparato di elaborazione dati 6 secondo la presente invenzione consente all'utente di indagare allo stesso tempo le relazioni tra i messaggi dell'insieme di riferimento 50 da vari punti di vista, ovvero condurre un'analisi

sistemica di tali messaggi in modo facile ed intuitivo, interagendo con l'interfaccia utente 10.

Poiché l'interfaccia utente 10 è configurata per mostrare l'insieme di riferimento dei messaggi 50 come aggregato dall'apparato di elaborazione dati 6 nelle varie dimensioni dell'array 9, l'utente può ottenere dall'interfaccia utente 10 informazioni su vari aspetti delle relazioni tra i messaggi dell'insieme di riferimento (ad esempio ora, hashtag, ecc.) interagendo con l'interfaccia utente 10, senza la necessità di eseguire complesse operazioni come query multiple ecc.

In questo modo, l'utente può identificare tendenze, hashtag chiave, utenti chiave, luoghi, nuovi argomenti, ecc. associati al contesto di riferimento al quale l'insieme di riferimento di messaggi 50 si attiene.

Inoltre, le precedenti fasi eseguite dall'apparato di elaborazione dati 6 possono essere eseguite quasi in tempo reale, ovvero a blocchi mentre lo scambio di messaggi pertinenti al contesto di riferimento è in corso. L'utente può quindi eseguire un'analisi sistemica di detti messaggi in tempo quasi reale, in modo da poter utilizzare i risultati dell'analisi per intraprendere azioni all'interno del contesto di riferimento, ad esempio migliorare l'efficacia della comunicazione, aumentare la portata della discussione, incrementare la partecipazione alla discussione di un certo gruppo di utenti, ecc.

## RIVENDICAZIONI

1. Un metodo per elaborare messaggi (5) scambiati su un social network (4) da una pluralità di dispositivi utente (2) connessi ad una rete di comunicazione (3) che supporta detto social network (4), detto metodo comprendendo, presso un apparato di elaborazione dati (6) cooperante con detto social network (4) attraverso detta rete di comunicazioni (3):
  - a) acquisire da detto social network (4) un insieme di riferimento di messaggi (50) riguardanti un contesto di riferimento scelto dall'utente;
  - b) applicare un'aggregazione multidimensionale a detto insieme di riferimento di messaggi (50), detta aggregazione multidimensionale comprendendo raggruppare detti messaggi di detto insieme di riferimento (50) in un array di gruppi (9) avente almeno due dimensioni, detto array (9) comprendendo, in ogni dimensione, un numero di gruppi che corrispondono ad un numero di valori di una rispettiva caratteristica del messaggio, in cui ciascun gruppo di detto numero di gruppi aggrega tutti i messaggi di detto insieme di riferimento (50) la cui detta rispettiva caratteristica di messaggio ha uno stesso valore di detto numero di valori, e
  - c) mostrare una interfaccia utente (10) configurata per mostrare detto insieme di riferimento di messaggi (50) come aggregati in almeno due dimensioni di detto array (9).
2. Il metodo secondo la rivendicazione 1, in cui alla fase a) detto insieme di riferimento di messaggi (50) è ottenuto applicando a detti messaggi (5) scambiati su detto social network (4) una regola di filtraggio basata su uno o più valori di almeno una caratteristica di messaggio di filtraggio di

detti messaggi (5), detta almeno una caratteristica di messaggio di filtraggio essendo scelta tra:

- luogo di origine di detti messaggi (5);
- data e/o intervallo di tempo di pubblicazione di detti messaggi (5);
- autore di detti messaggi (5);
- username e/o hashtag e/o parola chiave citata in detti messaggi (5);
- lingua di messaggi(5);
- azioni eseguite su messaggi (5);
- tipo di contenuto di messaggi (5);
- caratteristiche del contenuto di messaggi (5).

3. Il metodo secondo le rivendicazioni 1 o 2, in cui la fase a) è eseguita da un programma software per l'importazione attraverso l'invocazione di una application programming interface di detto social network (4).

4. Il metodo secondo una qualsiasi delle rivendicazioni precedenti, in cui alla fase b) detta caratteristica di messaggio è scelta tra:

- luogo di origine di detto insieme di riferimento di messaggi (50);
- data e/o intervallo di tempo di pubblicazione di detto insieme di riferimento di messaggi (50);
- autore di detto insieme di riferimento di messaggi (50);
- username e/o hashtag e/o parola chiave citata in detto insieme di riferimento di messaggi (50);
- lingua di detto insieme di riferimento di messaggi (50);
- azioni eseguite su detto insieme di riferimento di messaggi (50);
- tipo di contenuto di detto insieme di riferimento di messaggi (50);

- caratteristiche di contenuto di detto insieme di riferimento di messaggi (50).
5. Il metodo secondo una qualsiasi delle rivendicazioni precedenti, in cui durante detta fase b) detto numero di valori di detta rispettiva caratteristica di messaggio sono valori predefiniti decisi da un utente.
  6. Il metodo secondo la rivendicazione 5, in cui detta fase b) comprende anche applicare un'ulteriore aggregazione a detto insieme di riferimento di messaggi (50), detta ulteriore aggregazione multidimensionale comprendendo espandere detto array di gruppi (9) aggiungendo, in un'ulteriore dimensione, un ulteriore numero di gruppi che corrisponde ad un ulteriore numero di valori, o di cluster di valori, di una caratteristica emergente dei messaggi scoperta in detto insieme di riferimento di messaggi (50) da detto apparato di elaborazione dati (6), ogni gruppo di detto ulteriore numero di gruppi aggregando tutti i messaggi di detto insieme di riferimento (50) la cui detta caratteristica di messaggio emergente ha uno stesso valore di detto ulteriore numero di valori o un valore appartenente ad uno stesso cluster di un ulteriore cluster di valori.
  7. Il metodo secondo una qualsiasi delle precedenti rivendicazioni, in cui detta interfaccia utente (10) è configurata per fornire di default una rappresentazione grafica di detto insieme di riferimento di messaggi (50) come aggregato in almeno una dimensione di detto array (9), quando l'utente la avvia.
  8. Il metodo secondo la rivendicazione 7, in cui detta interfaccia utente (10) è configurata per essere navigata da detto utente in modo da fornire una

- rappresentazione grafica di detto insieme di riferimento di messaggi (50) come aggregati in almeno una ulteriore dimensione di detto array (9).
9. Un apparato di elaborazione dati (6) comprendente un processore configurato per eseguire le fasi del metodo secondo una qualsiasi delle rivendicazioni da 1 a 8.
  10. Un programma per elaboratore comprendente istruzioni che, quando eseguite da un elaboratore, inducono l'elaboratore ad eseguire le fasi del metodo secondo una qualsiasi delle rivendicazioni da 1 a 8.

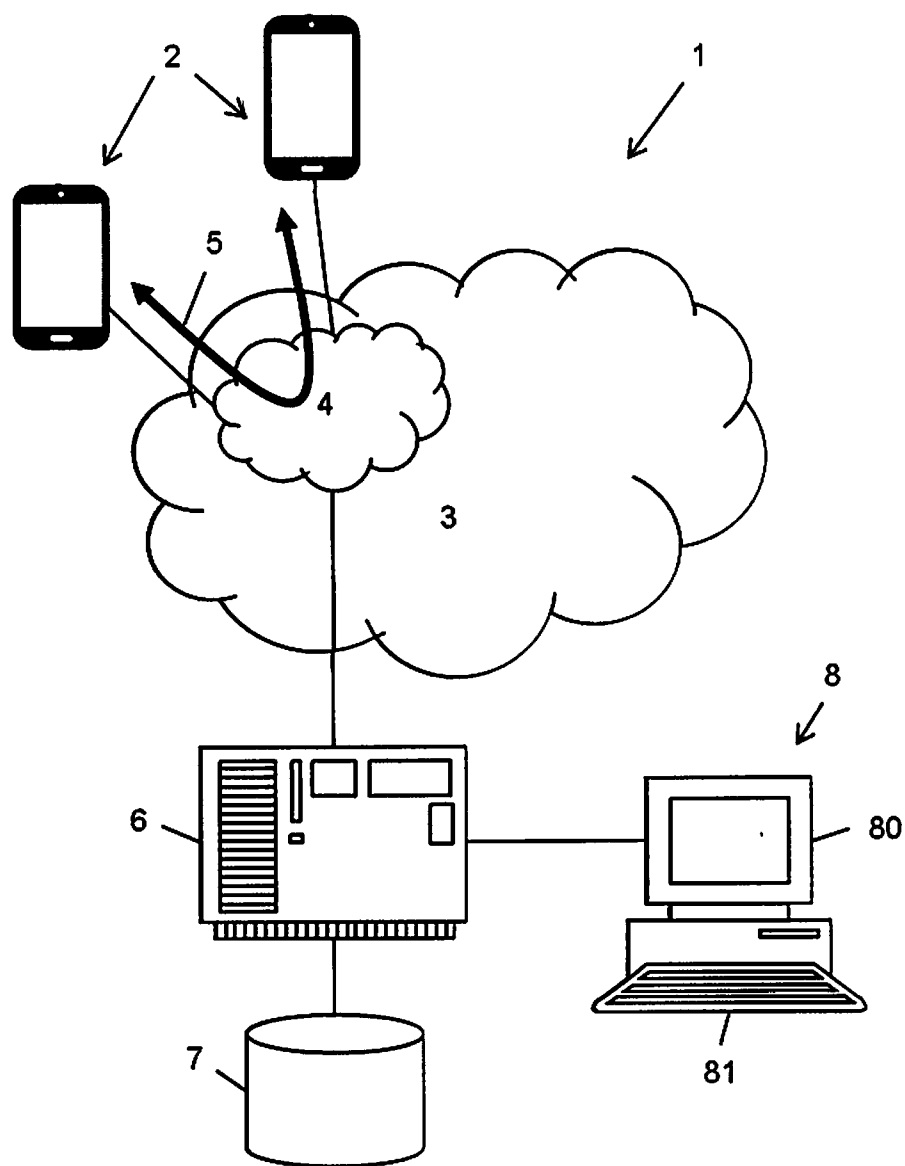


Fig. 1

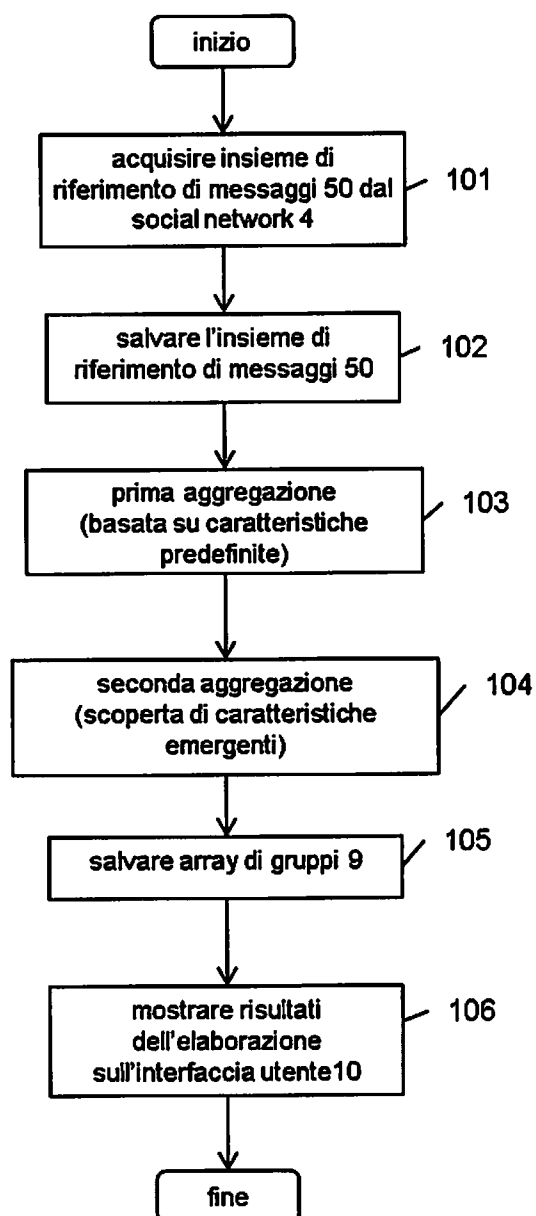
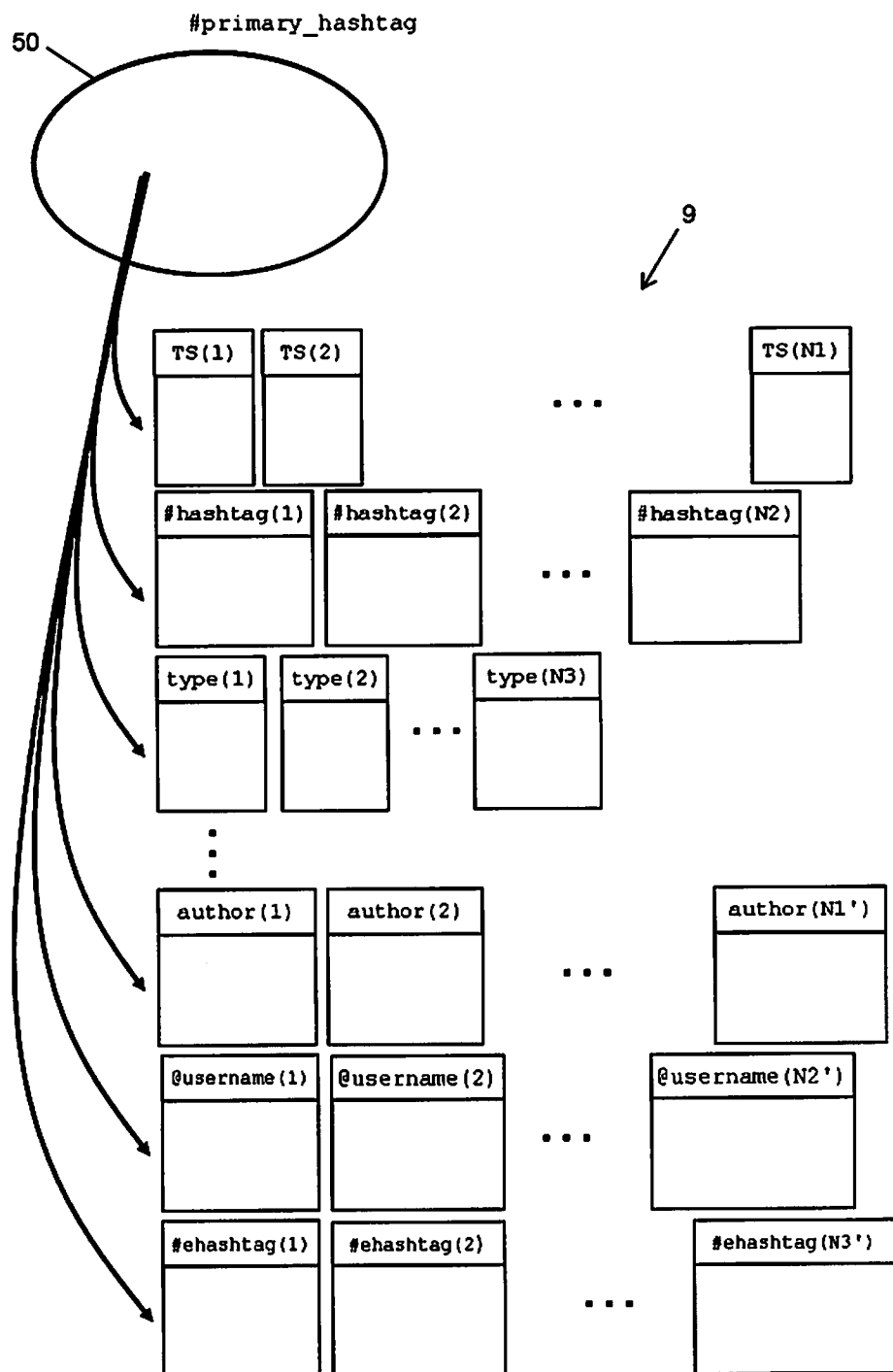


Fig. 2

Fig. 3

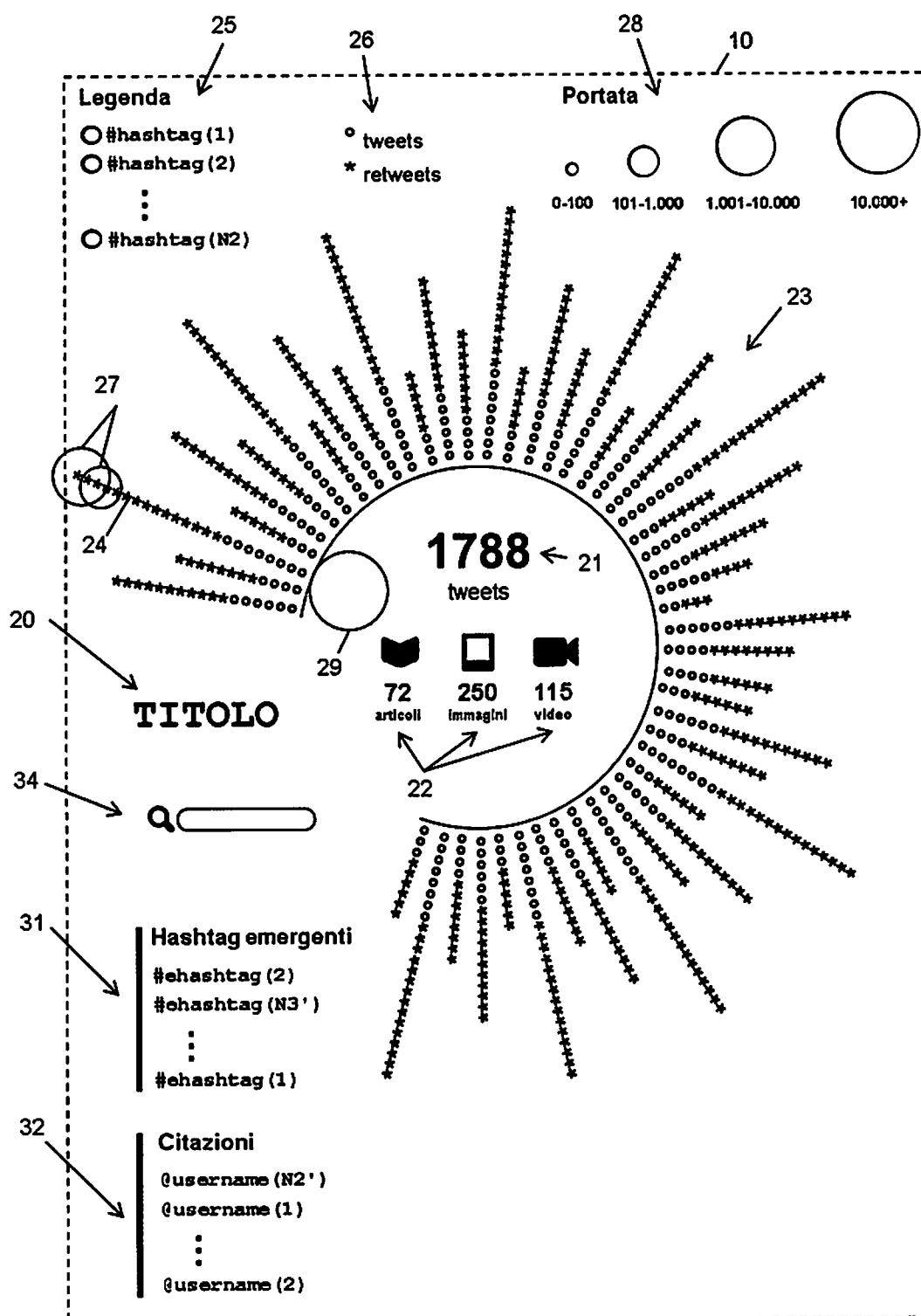


Fig. 4a

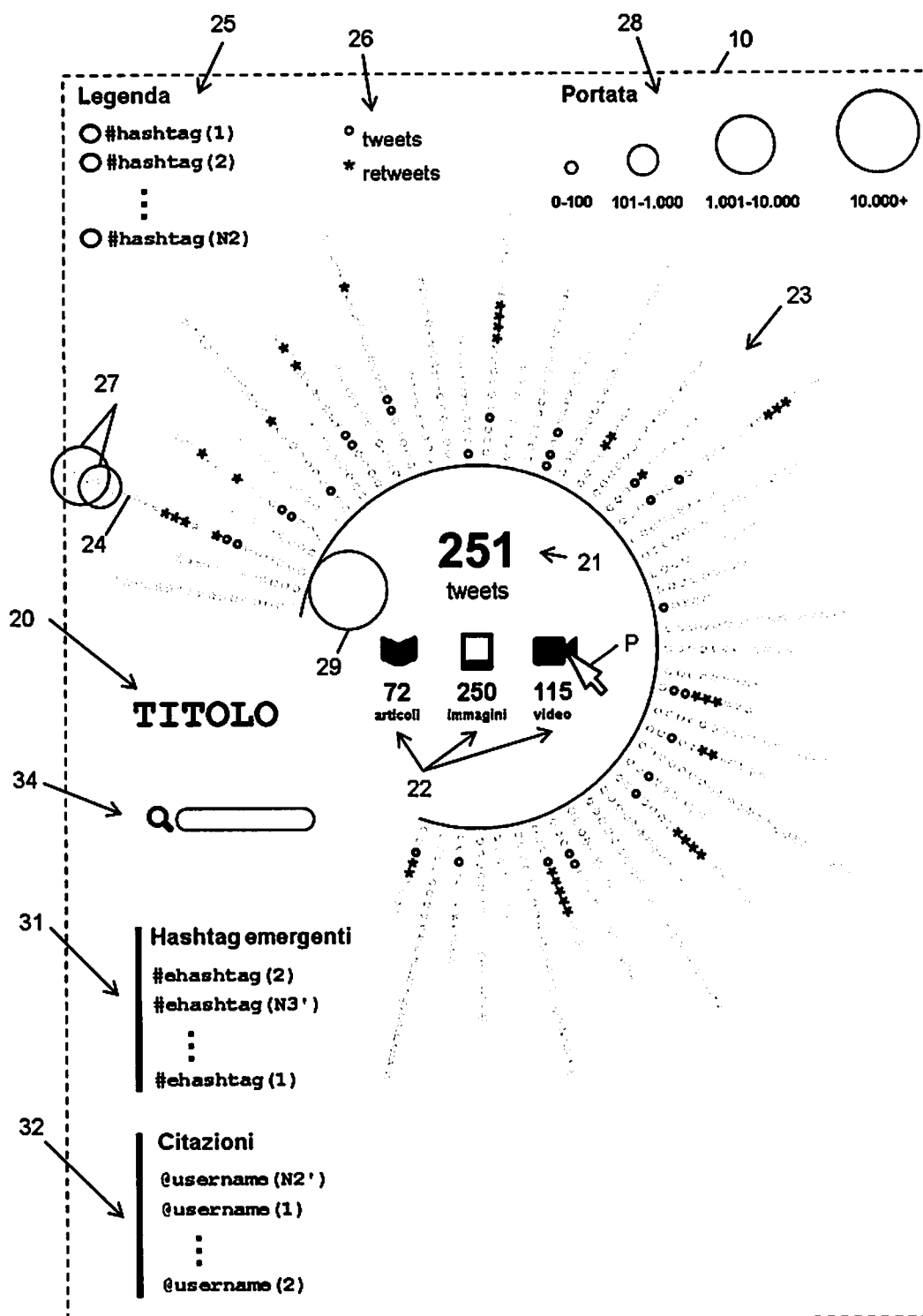


Fig. 4b

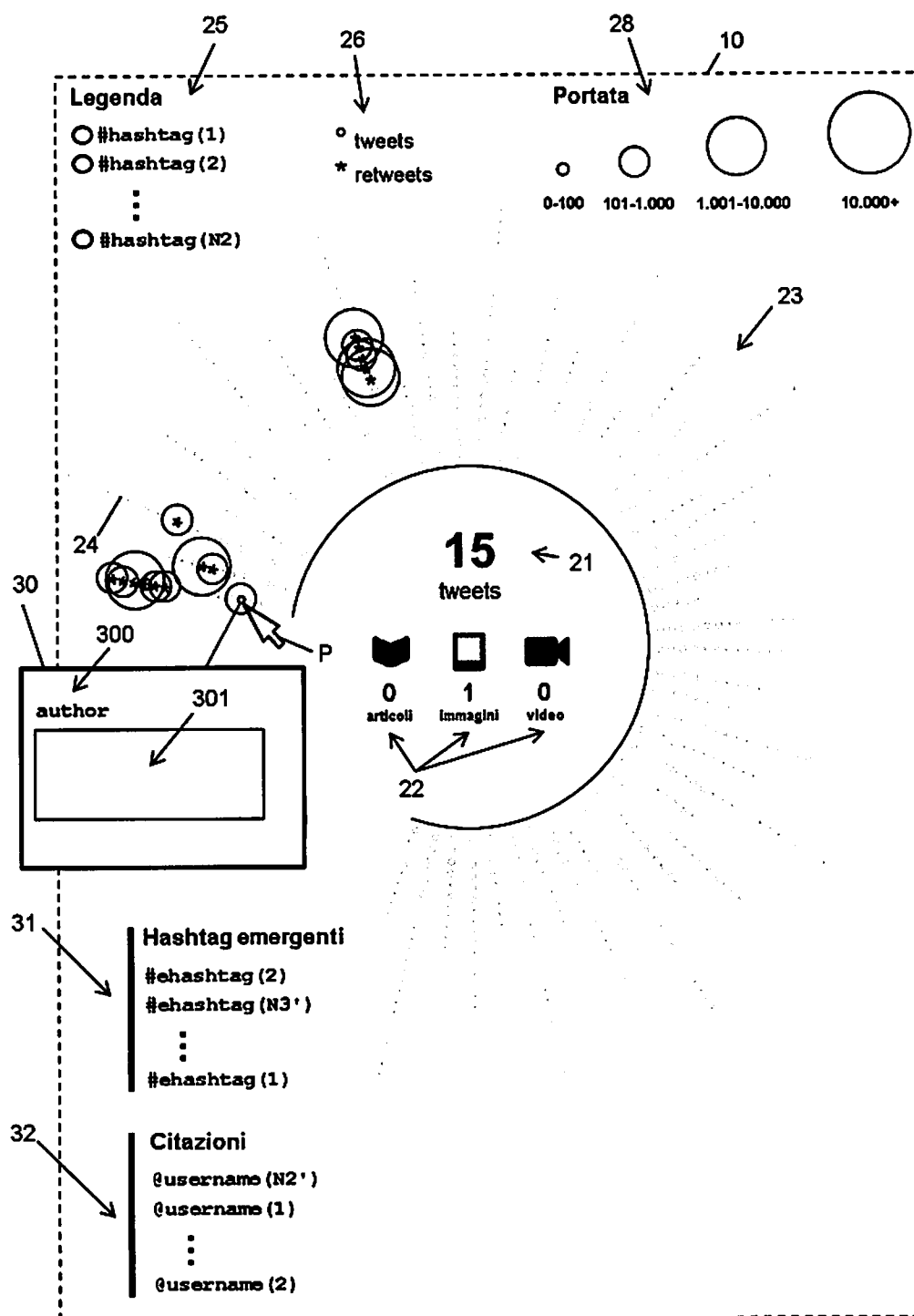


Fig. 4c

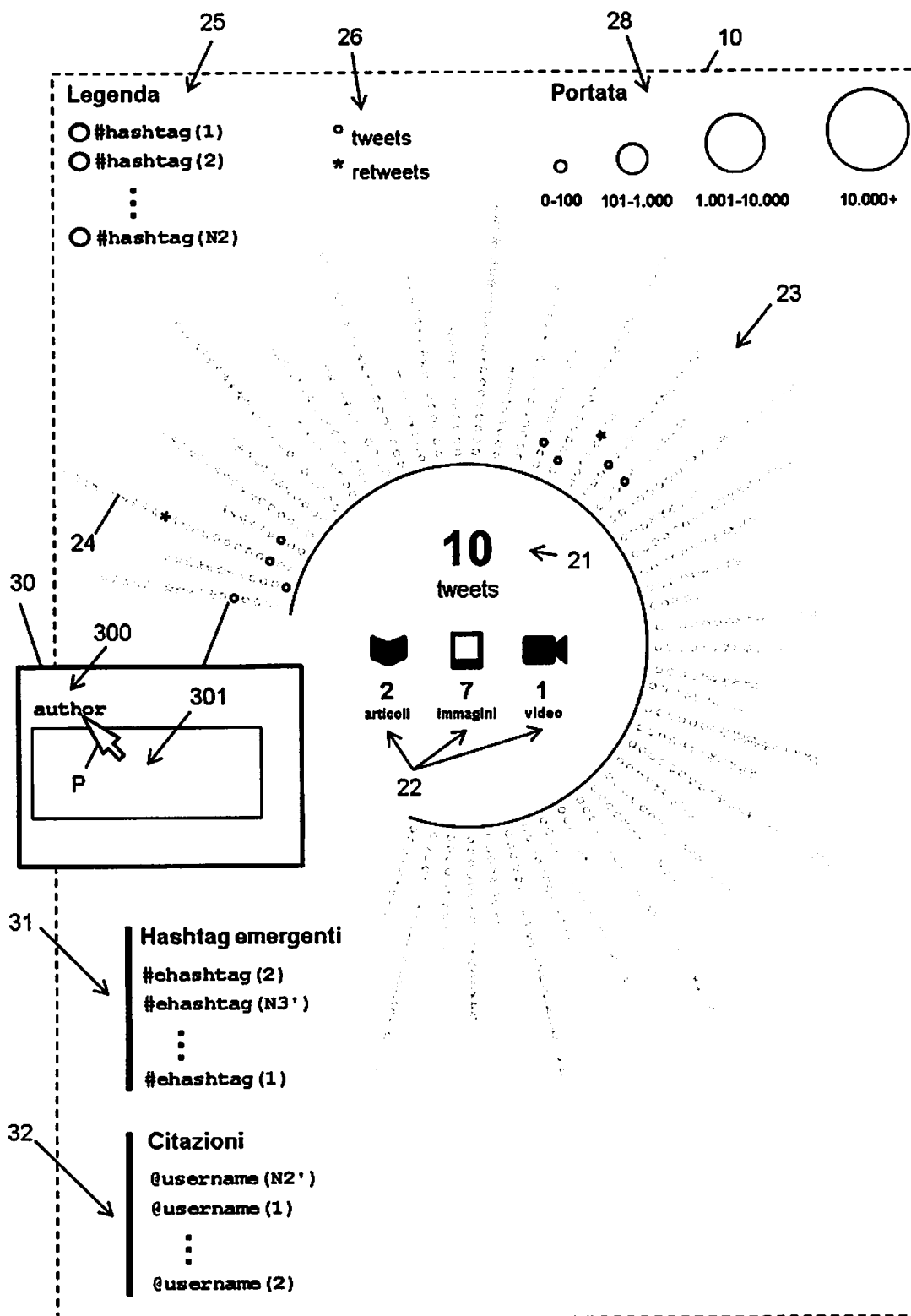


Fig. 4d

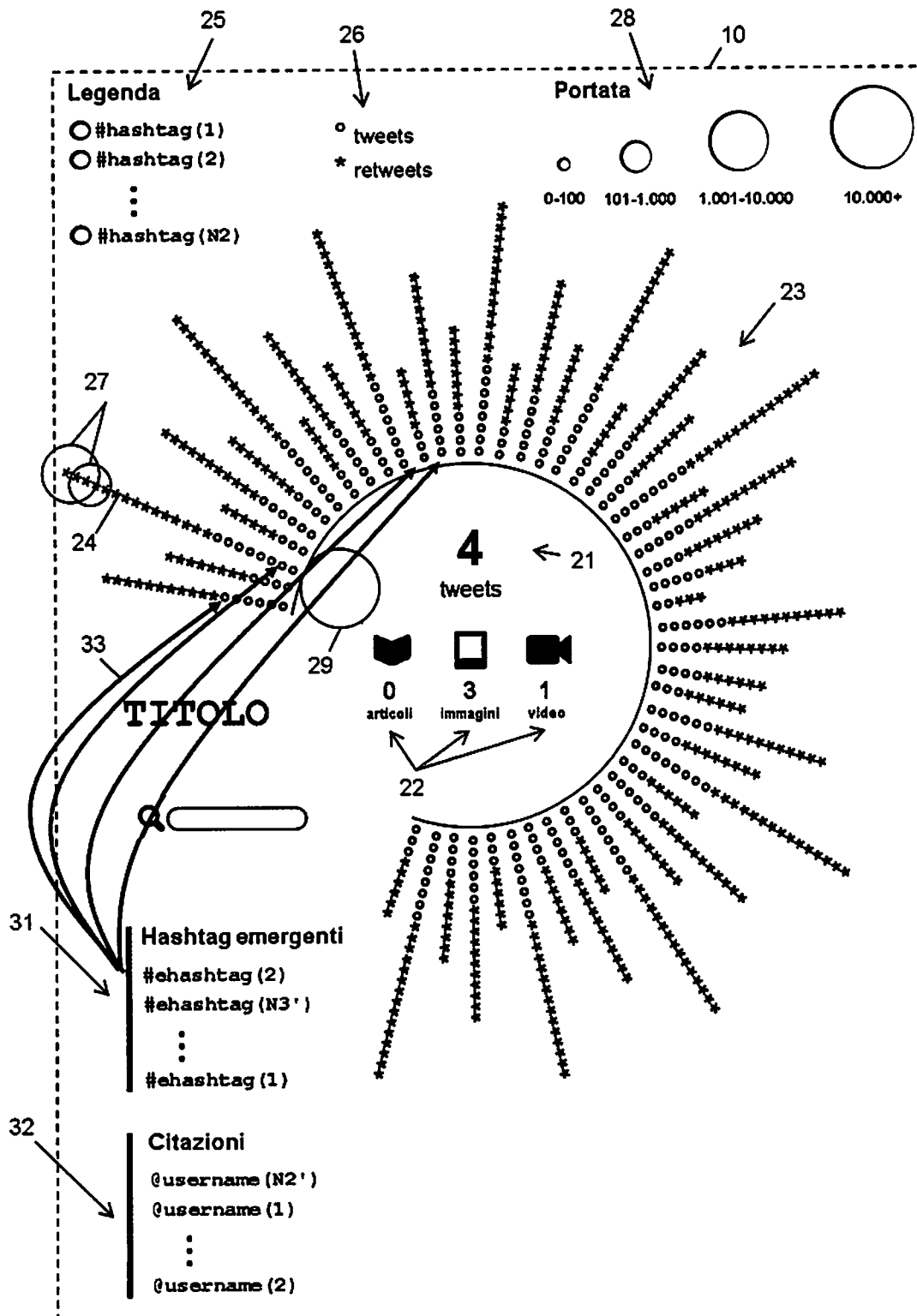


Fig. 4e