

SOCIAL SENSING

Dalla rete agli utenti
e ritorno

MAURIZIO TESCONI

Le attuali piattaforme dei Social media consentono ogni giorno a milioni di persone di condividere informazioni e documentare eventi in tempo reale. È possibile utilizzare gli utenti delle piattaforme Social come sensori da impiegare al verificarsi di eventi nel mondo reale? Nel caso di un disastro naturale, come di un attacco terroristico, sarebbe possibile richiedere agli utenti della rete presenti sul posto di scambiare nell'immediatezza informazioni sull'evento?

La casa trema, gli oggetti iniziano a cadere, ma la tentazione di scattare una foto e condividerla prima di mettersi in salvo è grande. Durante il terremoto dell'Emilia del 2012, solo pochi secondi dopo la scossa di magnitudo 5.9 qualcuno già postava su Twitter una foto realizzata in una zona vicina all'epicentro. Vi è chi parla di *Social Network addiction*, ovvero una sorta di dipendenza dal vivere perennemente connessi, aggiornando e controllando di continuo il proprio profilo.

LA PAURA CORRE SUI SOCIAL

L'uso compulsivo dei social network da parte di persone coinvolte in un'emergenza può diventare in verità una grande opportunità. Nelle situazioni di crisi gli utenti si comportano come veri e propri sensori, più propriamente definibili come 'sensori sociali'.

Negli ultimi anni i ricercatori hanno pensato di sfruttare questa enorme quantità di dati scambiati dagli utenti dei social media sotto forma di messaggi od oggetti multimediali.

In particolare, durante gli eventi improvvisi e spaventosi i dati possono fornire in tempo reale il quadro di quello che è accaduto.

Le situazioni di crisi che si possono generare a seguito di un disastro naturale o di un attentato terroristico spesso costringono chi deve gestire le emergenze a prendere decisioni sulla base di dati il più delle volte insufficienti e poco attendibili per la difficoltà di un riscontro immediato su quanto effettivamente si è verificato. Per questo, le informazioni scambiate dagli utenti possono anche risultare utili per decidere di indirizzare tempestivamente i soccorsi verso le località maggiormente interessate dall'evento.

Secondo gli esperti del settore, il fenomeno Social media nasce nel 2001, con Wikipedia, l'enciclopedia scritta in maniera collaborativa da milioni di persone in tutto il mondo.

Negli anni successivi l'uso dei social network ha visto una crescita esponenziale, coinvolgendo un numero sempre maggiore di persone, fino a esplodere nel 2003 con MySpace, una comunità virtuale che offre blog, profili personali, gruppi, foto, musica e video. Pochi anni più tardi sono emersi Facebook, Youtube e Twitter che rappresentano attualmente i social più diffusi e producono ogni giorno una quantità di dati talmente grande da aver imposto l'utilizzo di tecnologie ad hoc per la loro gestione. Questi dati hanno tutte le 3 V che caratterizzano i big data: volume, velocity, variety. Infatti, raggiungono dimensioni enormi, sono prodotti molto velocemente e in diversi formati.

Ci troviamo di fronte a una vera e propria miniera da sfruttare: Facebook conta attualmente 1.6 miliardi di utenti attivi giornalmente, su Twitter si producono più di 500 milioni di tweet al giorno e su Youtube, ogni minuto, vengono caricate più di 100 ore di video. Recentemente, è stata introdotta una quarta V (veridicity) che ha lo scopo di mettere in guardia sull'accuratezza delle informazioni. Per esempio, il fatto che Twitter non richieda il nome reale dell'utente e non compia verifiche dell'identità agevola la diffusione di profili falsi o anonimi.

La diffusione dei social media apre nuovi scenari di utilizzo anche in caso di catastrofi naturali. Già nel 2004, a seguito dello Tsunami nell'oceano Indiano, le persone coinvolte utilizzarono un forum durante l'emergenza, mantenuto attivo per i dieci giorni successivi all'evento. Analogamente, un numero significativo di messaggi furono

scambiati su Myspace nel 2005, durante l'uragano Katrina, mentre Twitter venne largamente usato durante gli incendi che esplosero vicino San Diego, in California, nel 2007. Da quel momento, è diventata una pratica abbastanza comune utilizzare questo strumento di microblogging per comunicare, condividere informazioni od organizzare i soccorsi in presenza di situazioni di emergenza. Durante lo Tsunami in Giappone del 2011, provocato da un terremoto di magnitudo 9.0, il sistema telefonico giapponese andò fuori uso e Twitter divenne il principale mezzo di comunicazione, facendo registrare volumi di 1200 tweet al minuto nella sola città di Tokio.

SOCIAL SENSING: GLI UTENTI COME SENSORI SOCIALI

Gli utenti, con i loro smartphone, si sono trasformati in veri e propri sensori, che riportano in maniera dettagliata ciò che vedono, grazie a descrizioni testuali dell'evento, immagini, coordinate Gps e altri metadati utili per farsi un'idea di quello che è avvenuto.

Con il termine *crowdsensing* ci si riferisce alla condivisione di dati raccolti da sensori con lo scopo di misurare un fenomeno d'interesse. Esistono due tipi di crowdsensing, quello partecipativo e quello opportunistico. Nel primo, gli utenti sono reclutati e conoscono già il tipo di messaggio che devono fornire a seguito di un evento. Un caso noto è quello del quotidiano britannico «The Guardian» che, nel 2011, ha chiesto la collaborazione dei suoi lettori per passare al setaccio le decine di migliaia di documenti relativi alle spese scandalo dei parlamentari del Regno Unito: il giornale mise a disposizione dei lettori un'interfaccia per annotare ed evidenziare le informazioni contenute nei documenti, nel tentativo di scoprire ulteriori casi di corruzione e di illeciti. In Italia esiste Emergenza24¹, community di oltre 10.000 persone, che invita i propri iscritti a riportare un evento codificando le informazioni all'interno di un tweet. Nel crowdsensing opportunistico, diversamente, gli utenti non sanno di svolgere la funzione di sensori e sta al sistema che analizza i dati distinguere le informazioni importanti dal rumore di fondo per fornire alle autorità competenti un quadro della situazione più accurato possibile. Quest'ultimo è quello più interessante dal punto di vista scientifico, perché pone di fronte a una serie innumerevole di sfide aperte.

1. <<http://www.emergenza24.org/>> [20 maggio 2016].

Ascoltando sui social alcune parole chiave relative alle emergenze, per esempio, è molto facile raccogliere messaggi che utilizzano quelle parole con accezioni diverse o non riferite a eventi in corso. Una delle prime sfide è proprio quella di distinguere questi messaggi e lasciare solo quelli prodotti dalle fonti primarie (cioè i testimoni oculari) utili a sintetizzare quello che è accaduto. Una volta filtrati i messaggi, è possibile fare la rilevazione di un evento con tecniche di *anomaly detection*, che vengono utilizzate per individuare una situazione di emergenza. È possibile ispezionare rapidamente i contenuti rilevanti delle discussioni degli utenti mediante *word cloud* (letteralmente, nuvola di parole) e tecniche di *topic discovery* che servono a enucleare le parole più frequentemente utilizzate oppure a individuare gli argomenti maggiormente discussi.

Una delle sfide principali da affrontare in questo ambito va sotto il nome di *automatic text summarization*, ovvero cercare di fare una sintesi, il più esaustiva e completa, della grande quantità di dati testuali raccolti. A seguito di un evento, infatti, vi è un picco di attività sui social con informazioni molto simili e ridondanti. In questa mole di segnalazioni diventa cruciale saper distinguere quelle che portano il maggior contenuto informativo e che meglio descrivono l'evento d'interesse. Un simile approccio può essere applicato alle immagini scambiate durante l'emergenza. In questo caso si utilizzano tecniche di *deep learning* per mettere insieme immagini simili e consentire all'utente di orientarsi, per esempio, organizzandole all'interno di un'interfaccia per navigare queste informazioni.

Negli ultimi anni, numerose iniziative di natura scientifica e applicativa sono state sviluppate nell'ambito del Social Sensing. Alcuni lavori presentati sono poi stati effettivamente implementati e risultano attualmente in uso, altri si sono invece fermati alla fase progettuale, senza dar seguito alle soluzioni proposte. Tra i sistemi per la gestione delle emergenze basati su Social Sensing sono stati recentemente proposti quelli di *smart public safety*, in grado di sfruttare sia reti di sensori wireless sia dati provenienti da social media. In tali contesti, i dati raccolti da sensori tradizionali sono arricchiti con informazioni provenienti da sensori sociali utilizzando metodologie di *data fusion*, dal momento che tali software devono poter gestire e analizzare dati di natura e di origine differente.

Tra i vari tipi di emergenze, in letteratura scientifica i terremoti sono stati molto studiati perché, tramite i sismografi, è possibile avere accesso a un ricco insieme di dati che possono essere sfruttati per lo sviluppo e la validazione di modelli statistici o di nuovi algoritmi.

In particolare, le opportunità di applicazione di metodologie di Social Sensing per la gestione di emergenze sismiche sono state descritte per prime in alcuni lavori², a partire dal 2010. Dando seguito a queste embrionali proposte, sono stati sviluppati e validati una serie di sistemi con caratteristiche simili, come Ears³, di supporto alle decisioni per la gestione delle situazioni di emergenza, sviluppato dal Cnr di Pisa. Essi raccolgono dati da social media e li analizzano in tempo reale. Lo scopo è di accorgersi dell'occorrenza di un evento sismico e di estrarre quante più informazioni possibili sull'accaduto, il tutto nel minor tempo possibile e in modo completamente automatico. Un approccio simile è stato adottato per la realizzazione della piattaforma Esa⁴ (Emergency Situation Awareness). In questo caso, le tecniche di Social Sensing non sono adoperate solo in ambito sismico, perché il sistema è in grado di segnalare emergenze di vario tipo, da alluvioni e frane, a incidenti aerei e rivolte in strada. La sua maggiore generalità sconta, tuttavia, una minore accuratezza dei risultati. Inoltre, una serie di altre iniziative ha visto l'applicazione di sistemi basati su Social Sensing per l'individuazione di incendi, code e ingorghi stradali o generici disservizi del trasporto pubblico. Recentemente, un gruppo di ricercatori ha realizzato uno studio estensivo della letteratura in questo ambito⁵, riferimento per ulteriori iniziative nell'ambito del Social Sensing, del Social Media Emergency Management (Smem), del Crisis Informatics e del Crisis Mapping.

LE PROSSIME SFIDE NEL CAMPO DEL SOCIAL SENSING

In futuro, la partita della ricerca in questo ambito si giocherà prevalentemente su due sponde: l'elaborazione automatica delle mappe di crisi, da un lato, e il coinvolgimento dei testimoni oculari dell'evento, dall'altro.

Immaginiamo l'impatto che una catastrofe di qualsiasi natura potrebbe avere su un ampio territorio. Disporre di una rapida panoramica geografica, a seguito dell'evento, consente di predisporre opportune contromisure e pianificare in modo opportuno le azioni successive. Prendiamo come esempio un attentato terroristico o un'alluvione. La conoscenza dell'accaduto in tempi brevi è importante, sia per allertare le autorità competenti (Forze dell'ordine o enti di protezione civile ecc.), sia per segnalare ai cittadini che alcune aree sono a rischio e che è necessario evitarle.

Le mappe di crisi sono mappe tematiche costruite utilizzando dati georiferiti per riportare, su una dimensione spaziale, i contenuti dei social media. Ciò si traduce, ad esempio, nella possibilità da parte degli enti di disporre in pochi minuti di una prima visione d'insieme dell'impatto di una catastrofe naturale su una vasta area popolata.

2. EARLE 2010, pp. 221-222; SAKAKI - M. OKAZAKI - Y. MATSUO 2010, pp. 851-860.

3. AVVENUTI ET AL. 2014, pp. 1749-1758.

4. <<http://www.esa.csiro.au/esa/welcome.html>> [20 maggio 2016].

5. IMRAN ET AL. 2015, p. 67.

Per fare questo, oltre all'informazione geografica, è importante poter disporre di sistemi di *machine learning* capaci di individuare e classificare i danni riportati dalle persone all'interno dei messaggi dei social network. Oltre alle informazioni che già nascono georiferite all'interno dei social media, si possono estrarre informazioni di natura geografica tramite il *geoparsing*, ovvero il riconoscimento automatico di luoghi all'interno dei testi condivisi dagli utenti. In zone densamente abitate, è possibile realizzare specifiche mappe della città che evidenziano le zone più colpite o a rischio. Poter utilizzare immediatamente queste informazioni, da chi è preposto all'organizzazione della prima risposta all'evento e del primo soccorso, è un innegabile vantaggio in quanto è possibile dislocare le risorse necessarie e in quel momento disponibili in maniera più razionale. La disponibilità di mappe interattive migliora la leggibilità dei dati in quanto si può aggiungere un'ulteriore dimensione permettendo, ad esempio, l'esplorazione dei messaggi delle zone ritenute più esposte in cerca di richieste di aiuto. Grazie a questo tipo di visualizzazione è possibile venire a conoscenza della mancanza di servizi essenziali o di danneggiamenti nelle zone evidenziate nella mappa. Tipicamente, il lavoro della protezione civile può richiedere alcune settimane per fornire un'adeguata stima dei danni, mentre con queste tecniche, già dopo poche ore, si possono ottenere alcune stime, forse meno precise ma che non si discostano significativamente dal dato definitivo.

Come accennato, i social offrono la possibilità di raccogliere informazioni utili anche coinvolgendo direttamente gli utenti. Il *participatory crowdsensing* è il paradigma che descrive questo scenario: gli utenti possono partecipare attivamente e consciamente alla raccolta di informazioni, trasmettendo, per esempio, attraverso i loro dispositivi, dati e rilevazioni utili su ciò che li circonda. Se da un lato l'approccio opportunistico si concentra più sul recupero di informazioni al volo ma, necessariamente più generiche, l'approccio partecipativo consente di raccogliere informazioni specifiche e puntuali, un compito che richiede tempo e varia a seconda della posizione del disastro. Ma chi sono effettivamente questi volontari digitali? Si può trattare di volontari digitali reclutati in anticipo o di testimoni oculari individuati in 'near real time' sulla base dei dati raccolti in fase opportunistica.

Un esempio del primo approccio è la piattaforma 'Did you feel it?'⁶ creata dalla Usgs, l'agenzia scientifica geologica degli Stati Uniti. Si

6. <<http://earthquake.usgs.gov/data/dyfi/>> [20 maggio 2016].

tratta di un questionario inviato agli utenti che hanno precedentemente condiviso la loro adesione al programma in caso di terremoto. Nel questionario, i cittadini coinvolti devono fornire una descrizione completa del fenomeno, compresi gli effetti del terremoto e l'entità del danno. Un esempio simile è la piattaforma di Crisis Mapping Ushahidi⁷, che consente di raccogliere, attraverso diversi canali tra cui social media, mail e sms, informazioni circa l'intensità di un disastro naturale.

Sul fronte dei social network, Facebook, con il servizio *Safety check*⁸, permette di avvisare la propria rete di amici sul proprio stato di salute e controllare la lista degli amici che potrebbero trovarsi nella situazione d'emergenza. Il sistema, studiato per aiutare ad avere notizie di persone che si trovano in zone colpite da calamità naturali, è stato sperimentato per la prima volta nel 2011 dopo lo tsunami in Giappone e viene oggi utilizzato anche in drammatici casi, come gli attentati di Parigi e gli attacchi di Bruxelles. Esistono poi tentativi di coinvolgere direttamente gli utenti, testimoni oculari di un determinato evento. In particolare, diversi esperimenti hanno l'obiettivo di scoprire le caratteristiche tipo del testimone oculare che riporta sui social informazioni su un dato accadimento. Il fine ultimo è di ingaggiarlo sul momento per chiedergli informazioni più dettagliate su ciò che sta vivendo.

Utilizzare i dati dei social media per fornire supporto nella gestione delle emergenze comporta anche l'esigenza di affrontare tutti i problemi relativi alla gestione della privacy degli utenti e le policy dei social network. La privacy all'interno dei social media è un diritto tutelato dalla legge e ha lo scopo di garantire che le informazioni sugli utenti siano trattate nel rispetto delle norme che le regolano e, in particolare, nel rispetto della volontà del singolo alla loro diffusione o meno. I dati condivisi possono essere considerati pubblici ma, per rispettare le policy imposte dai vari social media, è necessario salvarli solo temporaneamente e utilizzare tecniche di anonimizzazione che non consentano di ricostruire informazioni personali come, ad esempio, quelle relative agli spostamenti delle persone

G

7. <<https://www.ushahidi.com/>> [20 maggio 2016].

8. <<https://www.facebook.com/about/safetycheck/>> [20 maggio 2016].

BIBLIOGRAFIA MINIMA

- P. EARLE, *Earthquake Twitter*, «Nature Geoscience» 3 (2010), pp. 221-222.
 T. SAKAKI – M. OKAZAKI – Y. MATSUO, *Earthquake shakes twitter users: real-time event detection by social sensors*, in *Proceedings of the 19th international conference on world wide web*, «ACM Digital Library» (2010), pp. 851-860.
 M. AVVENUTI ET AL., *EARS (Earthquake Alert and Report System): a real time decision support system for earthquake crisis management*, *Proceedings of the 20th ACM SIGKDD international conference on knowledge discovery and data mining*, «ACM Digital Library» 2014, pp 1749-1758.
 M. IMRAN ET AL., *Processing Social Media Messages in Mass Emergency: A Survey*, «ACM Comput. Surv.» (CSUR), 4, 47 (giugno 2015).