DECIDE 2.0
Intelligent Processing of Citizen Opinions in Social Media

Carlos Iván Chesñevar

cic@cs.uns.edu.ar

http://cs.uns.edu.ar/~cic

CONICET / Laboratory of Research and Development in Artificial Intelligence

Department of Computer Science and Engineering
Universidad Nacional del Sur
Bahía Blanca, Argentina

W3C EGOV Presentation – August 24, 2012
Background

Government 2.0 refers to government’s adoption of Web 2.0 technologies to socialize government services, processes and data.

**Government 2.0 - benefits:**

- New ways of communication with citizens – e.g. through social media tools.
- New opportunities for government agencies to be informed about citizens’ needs and opinions through user-generated content.
Government 2.0 - challenges

The integration of data streams from social media into government 2.0 infrastructures poses several challenges:

- **Magnitude of information flow** – i.e. Twitter disseminates 55M tweets p/day; forces to rely on text mining (TM) and opinion mining (OM) techniques to filter noise and detect topics of community discussion.

- **TM and OM techniques are not common practices in government.**
Government 2.0 - challenges (cont.)

- Social media data streams are usually incomplete or potentially inconsistent, as citizens might have different views on a certain issue.
- Citizens’ arguments in social media must be assessed and confronted by government officials in order to be used as inputs in government decision making processes.
- To build trust, some decisions made by government need to be backed by arguments when informed to citizens.
About DECIDE 2.0

- **DECIDE 2.0**: a LACCIR research project supported by Microsoft Research, IDB and CONACyT (Mexico).
- **Research team**: Argentinean and Mexican researchers; Elsa Estevez (external consultant, UNU-IIST).

- **Research Problem**: To combine context-based search and argumentation in a collaborative system for managing service- and policy-related information in social media tools used by governments.

- **Project Aim**: a framework for intelligent processing of citizens’ opinions in social media, based on a collaborative system operating on top of existing social networks.
DECIDE 2.0 - goals (1/2)

- To implement models of trust and reputation propagation – users post information on social media whose reliability has to be assessed in order to effectively use such information for decision making.
- To develop algorithms for integrating information coming from different sources – several users may post messages related to the same topic; accrual of information needs to be modeled properly.
DECIDE 2.0 - goals (2/2)

- To design effective context representations and community identification algorithms — when analyzing citizen opinions, emerging communities have to be identified, and associated contextual information is to be obtained.

- To develop customized information models — providing targeted information to various categories of stakeholders requires having different “views” of the issues under analysis.
DECIDE 2.0

Context C of e-gov related issues

Argument Computation from Opinions on C

Opinion extraction on C using context-based information retrieval

Argument-based decision making

Production of global assessment of citizens’ opinion (identifying pro and con arguments about C)

CITIZENS’ OPINION DATABASE

Citizens’ opinion on C (with details for policy-making oriented decisions)

Citizens post opinions

Social media tools

Government Officials post policy-related issues

facebook
twitter Windows Live
Brief discussion on ongoing work


- Set of Tweets for a given Query or Context (for the sake of ex.: #Greece)
- Each Tweet t contains a set of terms \( \{t_1, t_2, \ldots, t_k\} \)
- A mapping M that maps Tweets in a set of possible sentiments

Neutral tweet on C
Positive tweet on C
Negative tweet on C

http://tweetsentiments.com/analyze
Twitter-based argumentation framework

- Criterion Cr
- Aggregation Operator
- Set of sentiments
- All possible Twitter-based arguments
- Attack relation between TB-Arguments (e.g. positive and neutral arguments are in conflict)
A Twitter-based (TB) argument

- **Query Q**
  (context)

- **Prevailing sentiment**
  (pos, neg, neutral)

- **Set of Tweets**
  tweets returned by

**Intended meaning**: the citizens’ overall opinion on query “#Greece” is neutral according to criterion Cr.
From a given query $Q$ \{#Greece\} we can derive more specific queries, like $Q_1$ \{#Greece, bailout\} or $Q_2$ \{#Greece, money\}.

$Q$ subsumes $Q_1$ and $Q_2$.
Attacks between TB-arguments

Universe of all Tweets

#Greece
money

#Greece
Opinion Trees

Opinion on Q (root node)

Counter-Opinion

Counter-Opinion

Universe of all Tweets

Opinion trees can be recursively defined.
Algorithm GetOpinionTree

• **Input:** Query Q
  Agg Operator
  Criterion Cr

• **Output:** Opinion Tree OT rooted in Q with Agg under Cr

“#Greece”
Twitter API
All tweets from 21.7.2012 between 15.00 – 23.00, GMT
High-level Algorithm
GetOpinionTree

• IF length(Q) <= 140
  THEN Let \( \langle \text{Agg, Q, Sent} \rangle \) be the root node
  where \( \text{Arg} = \text{Agg}(Q,C) \) and Sent is \( s(\text{Agg}(Q,C)) \)
  IF there are other Hashtags or keywords in \( \text{Agg}(Q,C) \) that expand Q
  THEN Compute \( L = (h_1, h_2, \ldots) \) List of Hashtags and keywords that expand Q in \( \text{Agg}(Q, C) \) according to some threshold (for example frequency)
  FOR EVERY \( h_i \in L \) do
    Put GetOpinionTree(Q, \( h_i \), Agg, C) as subtree rooted in \( \langle \text{Arg, Q, Sent} \rangle \)

A Java prototype was developed. Empirical analysis is being carried out for different sample contexts.
Some recent publications


"Integrating Argumentation Technologies and Context-Based Search for Intelligent Processing of Citizens’ Opinion in Social Media" (C. Chesñevar, A. Maguitman, E. Estévez and R. Brena).


Conclusions and Future Work

- We contend that our approach can lead directly to improved coverage, scalability and context-awareness with respect to the current model of information delivery and retrieval in social networks.

- Governments can greatly benefit from the proposed solution by a) having adequate mining techniques to retrieve valuable information provided by citizens on social media, and b) by targeting different announcements to the appropriate group of government stakeholders.

- Current results using Twitter show that the underlying machinery for DECIDE 2.0 is indeed feasible, lending itself to software implementation. Formal properties are currently being studied.
Thanks for your attention!
Questions…?