



BIG DATA:
REVOLUTIONIZING SOCIAL SCIENCE
METHODOLOGY
DIPAK K. GUPTA

ROOT CAUSES OF STRIFE: THE NECESSARY CONDITION

A collective movement takes place only when there is a shared set of grievance: economic, political, ethnic, religious, or environmental

However, these factors of frustration and anger are but the necessary conditions for actual formation of collective movements



THE SUFFICIENT CONDITION

For sufficient condition, we need conduits which can channelize this feeling of shared frustration in an organized form

This is done by “Political Entrepreneurs,” the leaders of political parties, labor and student unions, leaders of NGOs, heads of religious orders



WHAT DO POLITICAL ENTREPRENEURS DO?

They provide a psychological “frame” through which events can be understood by their followers. This framing is done by:

Defining the boundaries of the aggrieved community (“Us” factor or the “In-group”)

Defining the offending group (“Them” or the “Out-group”)

Articulating a clear threat from “them” to “Us”)



A STRIFE TAKE PLACE WHEN

Along with the framing, the leaders of the organization call for a specific action of protest

these calls for action can specify a time, place, and mode of protest



TRIGGER EVENTS

Long simmering grievances pave the road for political movements, however, their beginnings require trigger events, such as an inflammatory action by the government, a sudden increase in prices, the enactment of an unpopular policy, sudden exposure of corruption, incidents of police brutality, a change in government through legal (illegal) or illegal (coup-de-etats) etc.



GOING VIRAL

Mass movements from the earliest times have been caused by rumors. The paths of those rumors must have traversed similar paths (without the technology) as they do in today's world.

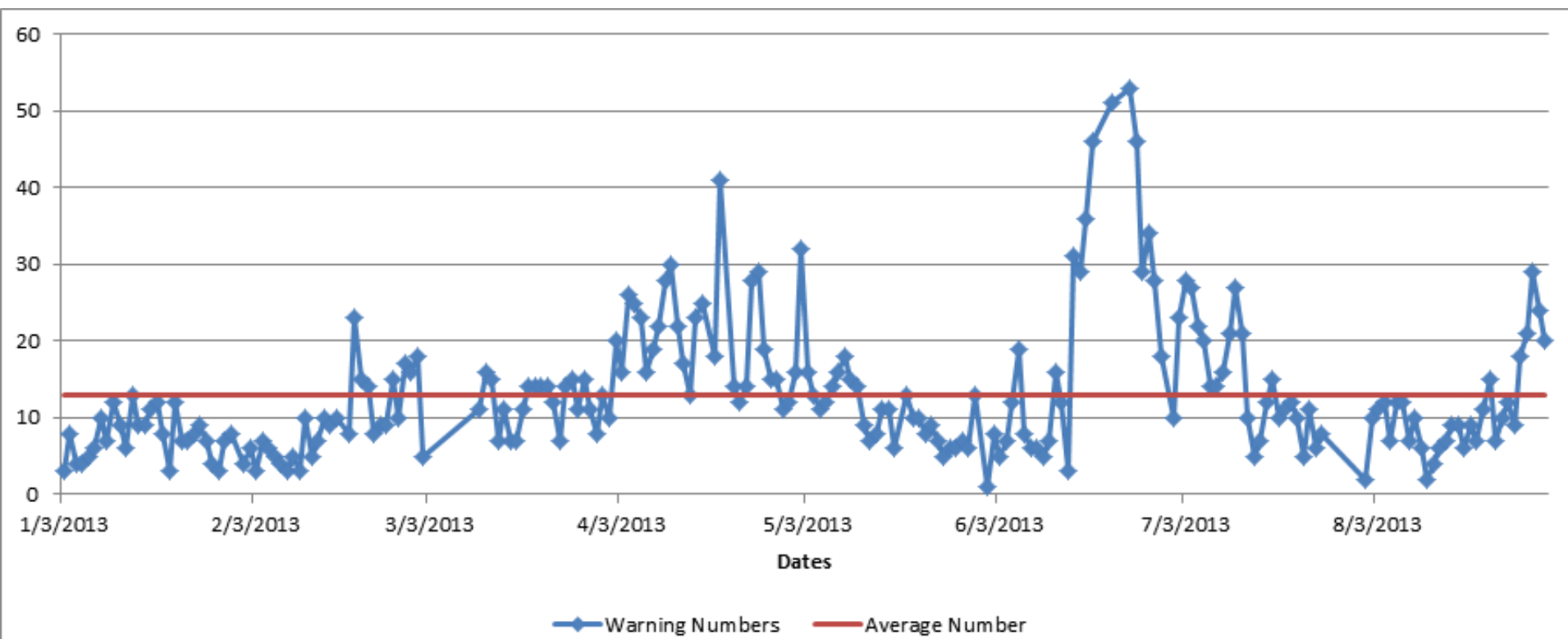
1. The best reason for a message going viral is the existence of repeatable visual or audio tape, which can be reproduced many times.
2. The message must be shocking (e.g., the picture of an actual beating or rape as opposed to the general knowledge that such things happen).
3. Ready identification with the individual victims (facial/voice recognition)
4. The "political entrepreneurs" repeating and rebroadcasting the messages in the social media
5. The messages jumping from social media to the traditional media
6. The messages entering into regional/national/international political discourse through the framing of "Us", "Them," and "Threat."

THE EMBERS EFFORT

So far, by using various models, we have been able to capture from the Internet, the words that would signify the necessary and sufficient conditions of civil strife. These include:



Ten Latin American Countries



FORECASTS OF CIVIL STRIFES FOR JULY 2013

Countries	GSR Events	# of Warnings	# Matches	Lead time
All countries	895	404	378	2.17
Brazil	269	132	132	2.09
Colombia	51	32	31	3.19
Uruguay	36	42	26	3.27
El Salvador	38	12	12	1.83
Mexico	166	41	41	1.34
Chile	32	29	26	2.35
Paraguay	107	28	28	1.75
Argentina	71	17	17	1.75
Venezuela	112	53	53	2.15
Ecuador	13	18	12	2.45

THE LIMITS OF FORECAST

	Timing	
	Timing known	Unknown Timing
Causality Known	<i>Known Knowns</i> (almost certain events: preplanned demonstrations, protest occurring of specific, politically significant days)	<i>Known unknowns</i> * (Widespread grievance but unknown trigger events – cutting down trees in a park near Taksim Square in Istanbul)
Causality unknown	<i>Unknown knowns</i> (Regular events in social strife: perhaps the commemoration of emotionally charged events on specific days of religious/political significance)	<i>Unknown unknowns</i> (Black swan events: perhaps precipitated by natural disasters, political assassinations, or sudden death of a leader)