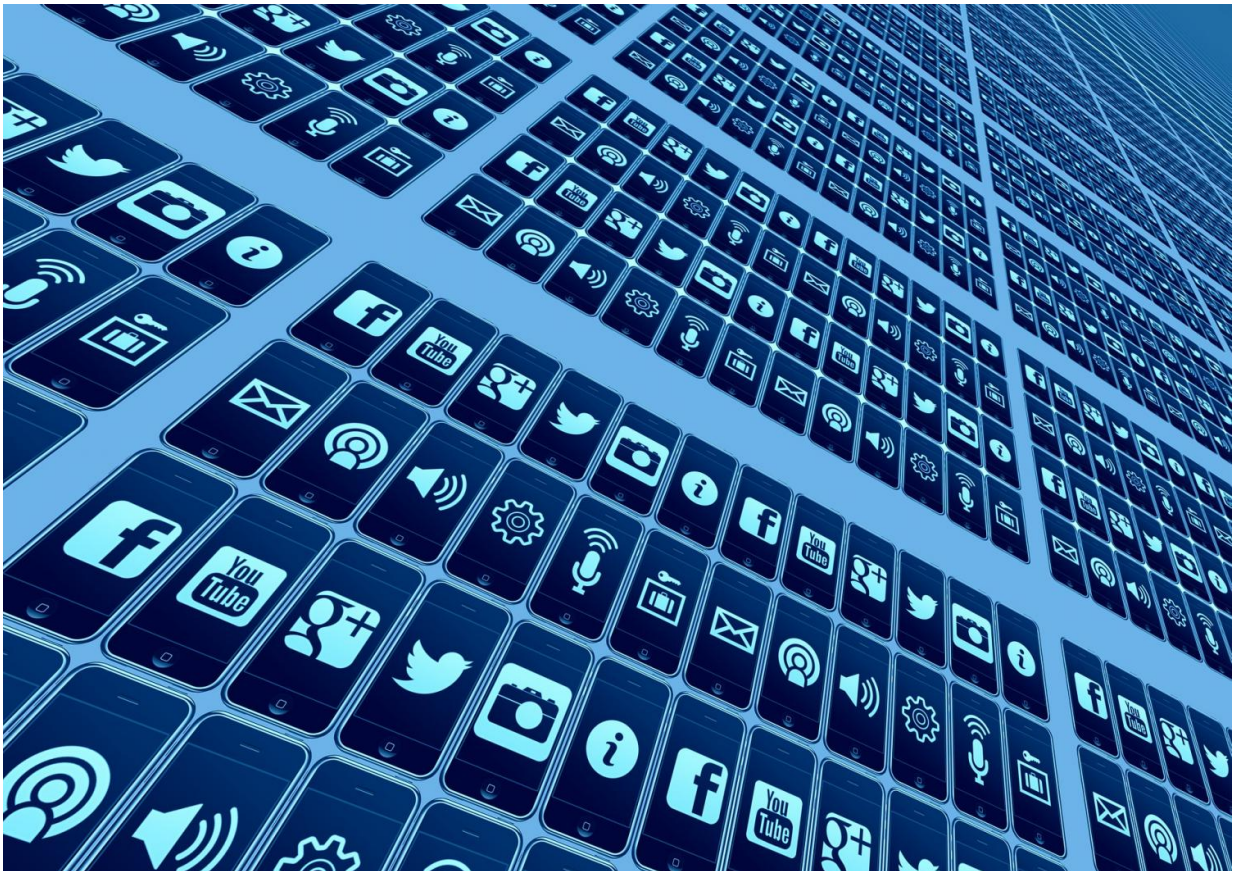


# Model fuses social media, remote sensing data with goal of identifying nuclear threats

July 23 2018

---



Credit: CC0 Public Domain

A new computational model allows researchers to draw on normally incompatible data sets, such as satellite imagery and social media posts,

to answer questions about what is happening in targeted locations. The researchers developed the model to serve as a tool for identifying violations of nuclear nonproliferation agreements.

"Our goal was to develop a working framework that uses information from a variety of sensors and [data](#) sources to identify these potential violations of nuclear nonproliferation," says Hamid Krim, co-author of a paper on the work, a professor of electrical and computer engineering at North Carolina State University and director of the VISSTA Laboratory. "Some of these data may be conventional, such as Geiger counter readings or multispectral data from [satellite imagery](#). But many of these data sources may be nontraditional, such as [social media posts](#). And these sources provide a wide variety of data that are not normally compatible, such as the text included on Twitter posts and the images posted on Flickr.

"By making these different inputs compatible with each other, we are able to accept a broader range of data inputs and use that data in a meaningful way that, ultimately, can help authorities reach more reliable conclusions," Krim says.

The researchers say the model can be used to work with any data that can be identified as coming from the targeted area. For example, satellite images are clearly identifiable, but they may also draw on social media posts that are actively or passively tagged as coming from the relevant area.

The question then becomes: how do you work with incompatible data? To explain, we'll use a proxy problem that the researchers used in their paper: identifying a flood. They chose a flood because data on flooding is not classified, whereas data regarding nuclear activity is.

The first step in the process is to use mathematical equations to translate

each type of data into a useful format. For example, images may be run through models to determine whether they are images of flooding, whereas text posts may be run through models to determine whether they include references to flooding. Once those data streams are translated into a neutral format – meaning they indicate flooding or no flooding – they can be compared to each other to answer basic questions such as: do the data support each other?

But it's not quite that simple. For example, people may be tweeting about a flood that is taking place hundreds of miles away, which could skew any calculation by the overarching model. To address this, the researchers incorporated mathematical elements that account for the complexity of the data they are drawing on.

"Addressing complexity is particularly important in the context of nonproliferation enforcement," Krim says. "Relevant data inputs may include photos of particular types of technology, references made in conversations caught on audio, and so on. A model like the one we developed needs to be flexible enough to account for the variability and complexity of both varied types of data and the varied clues we are looking for."

The researchers tested their [model](#) using data from a 2013 flood that took place in Colorado, and were able to resolve the incompatibility of multi-modal data in order to accurately estimate the location of the flooding.

Next steps for the project include evaluating nuclear facilities in the West to identify common characteristics that may also be applicable to facilities in more isolated societies, such as North Korea.

"We want to find ways of transferring information from known environment to a hidden one," Krim says. "How can we determine what

information and which models are transferable from one place to another, given incompatible or inconsistent data? What's normal, and what's not? It's not an easy problem."

**More information:** Han Wang, Erik Skau and Hamid Krim, "Fusing Heterogeneous Data: A Case for Remote Sensing and Social Media" *IEEE Transactions on Geoscience and Remote Sensing* (2018). [DOI: 10.1109/TGRS.2018.2846199](https://doi.org/10.1109/TGRS.2018.2846199)

## Abstract

Data heterogeneity can pose a great challenge to process and systematically fuse low-level data from different modalities with no recourse to heuristics and manual adjustments and refinements. In this paper, a new methodology is introduced for the fusion of measured data for detecting and predicting weather-driven natural hazards. The proposed research introduces a robust theoretical and algorithmic framework for the fusion of heterogeneous data in near real time. We establish a flexible information-based fusion framework with a target optimality criterion of choice, which for illustration, is specialized to a maximum entropy principle and a least effort principle for semisupervised learning with noisy labels. We develop a methodology to account for multimodality data and a solution for addressing inherent sensor limitations. In our case study of interest, namely, that of flood density estimation, we further show that by fusing remote sensing and social media data, we can develop well founded and actionable flood maps. This capability is valuable in situations where environmental hazards, such as hurricanes or severe weather, affect very large areas. Relative to the state of the art working with such data, our proposed information-theoretic solution is principled and systematic, while offering a joint exploitation of any set of heterogeneous sensor modalities with minimally assuming priors. This flexibility is coupled with the ability to quantitatively and clearly state the fusion principles

with very reasonable computational costs. The proposed method is tested and substantiated with the multimodality data of a 2013 Boulder Colorado flood event.

Citation: Model fuses social media, remote sensing data with goal of identifying nuclear threats (2018, July 23) retrieved 6 May 2024 from <https://techxplore.com/news/2018-07-fuses-social-media-remote-goal.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.