ADDRESSING GANG VIOLENCE IN THE DIGITAL AGE
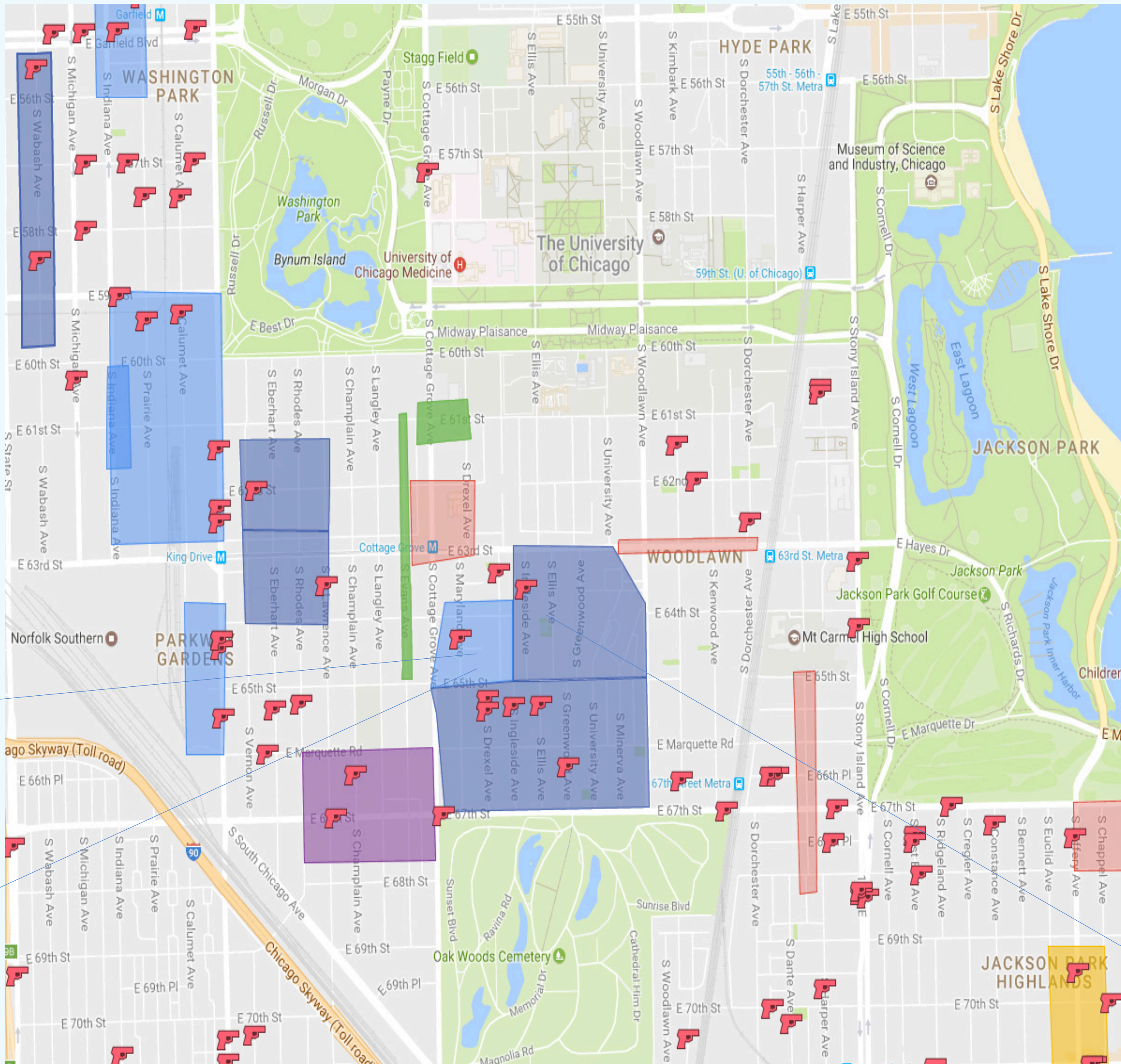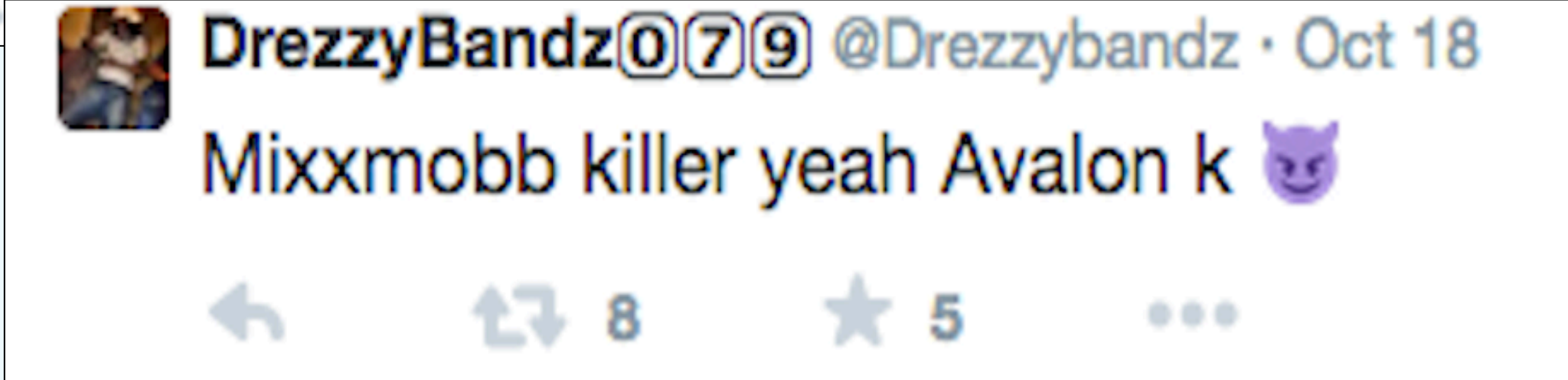## Using Twitter to Predict, Analyze, and Prevent Shootings in Chicago

Forrest Stuart • Alicia Riley

Overview

Chicago gang violence is at its highest levels in over a decade. Police, public officials, researchers, and community organizations agree that this upswing is related to (and potentially propelled by) gang feuds that occurring on social media platforms such as Twitter, Facebook, and Instagram. Using original datasets, this multi-method project will identify the association and potential predictive relationship between gang activity on Twitter and real-life violence. The ultimate aim of this research is to build an automated "early warning system" to help reduce and prevent gang violence.

> **LongLiveJuelz** @HotHeadTay_ · Nov 27
> Opps watching my every tweet 😂😂😂

> **DrezzyBandz0⑦⑨** @Drezzybandz · Oct 18
> Mixxmobb killer yeah Avalon k 😈
> ⟲ 8   ★ 5

> **DNT GO CMACK** @071bdrick · Nov 20
> FUCK SIRCONN STONYSPOT N 51ST YOUNG MONEY
> ★ 1

Next Steps

In the coming months, we will employ statistical techniques and unsupervised machine learning to pursue two primary questions:

Question 1: Are patterns in confrontational tweets associated with the timing and frequency of shooting incidents in/around gang territory? If so, which types of tweets have the strongest association.

Question 2: Which features of confrontational tweets precede and *predict* shootings in/around gang territory?
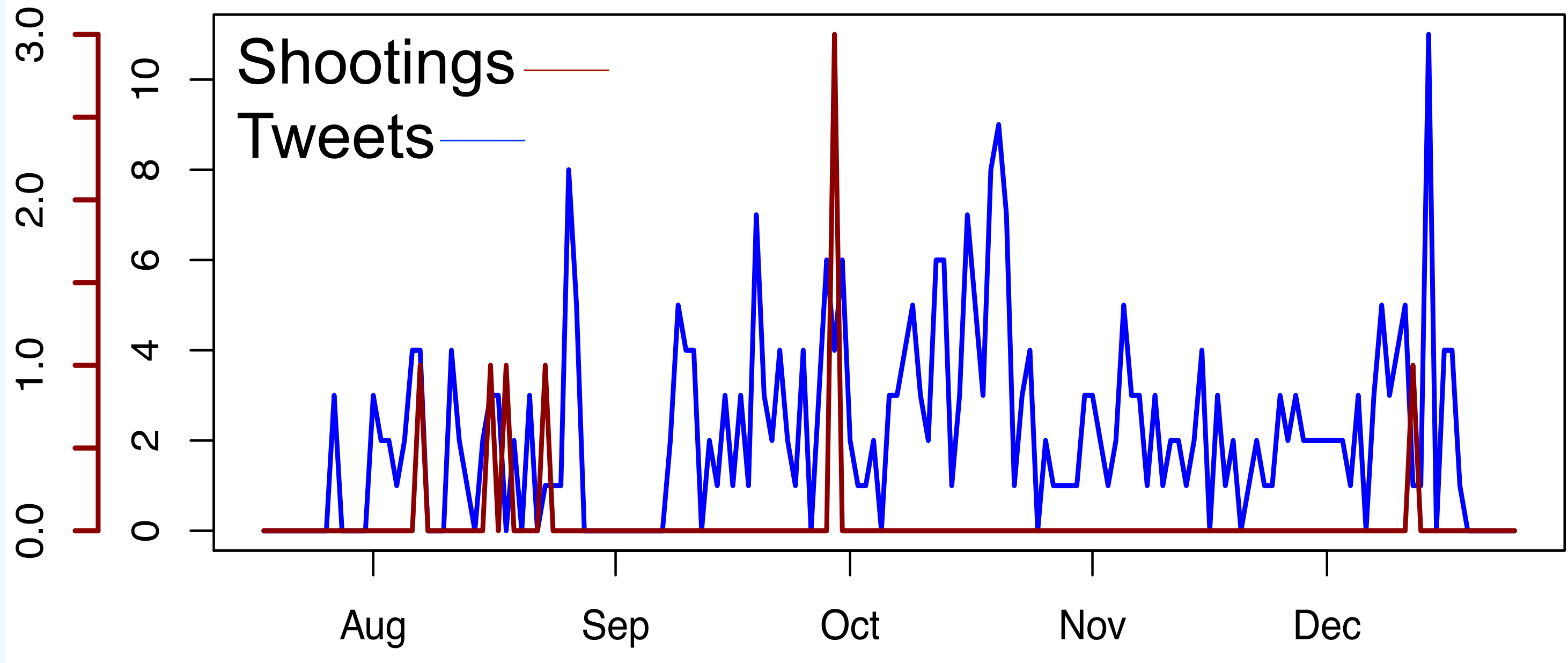
Constructing the Databases

Phase 1: Identifying Confrontational Tweets

Over two years of fieldwork alongside a Chicago gang, and 80 interviews with community residents, the research team catalogued specific Twitter keywords that surrounded shootings and gang feuds.

Phase 2: Searching the Twitter Stream for Confrontational Tweets

The research team built a program that searches the Twitter Rest API at regular intervals to collect all tweets that meet the characteristics identified in Phase 1. The database currently contains over 600,000 tweets.

### Phase 3: Building Data on Offline Violence

Using Chicago Police Department crime data, along with Geographic Information System (GIS) software, we identify all gun homicides and shootings in and around 57 different Chicago gangs.

Cleaning the Twitter Data Using Machine Learning

With the help of RCC collaborators, the research team is utilizing supervised machine learning to eliminate "noise" in the data. We have been training algorithms to classify tweets as relevant or not, informed by our field research. Only tweets classified as relevant are used to model associations with shootings.
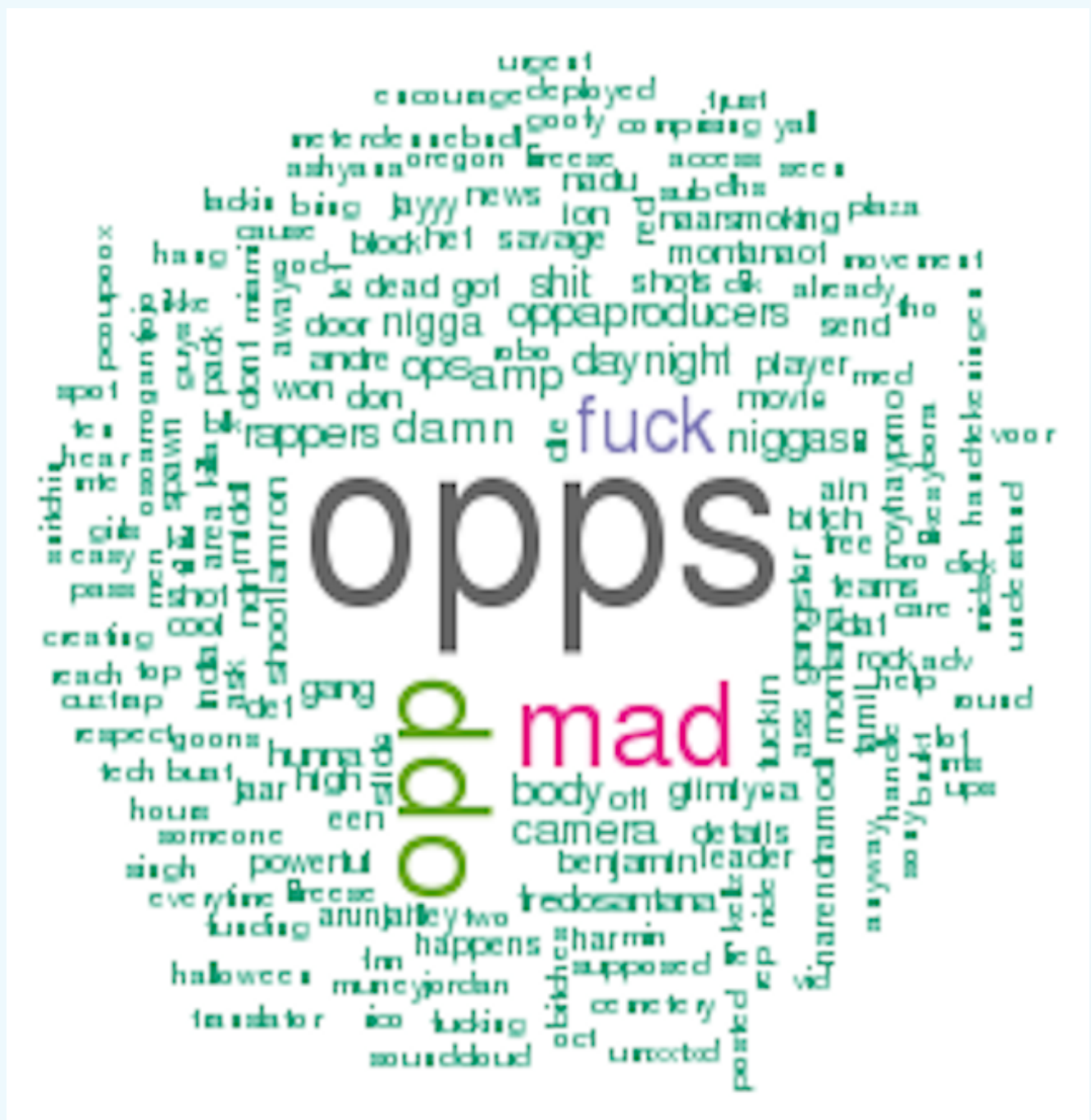
Word Cloud of Relevant Tweets
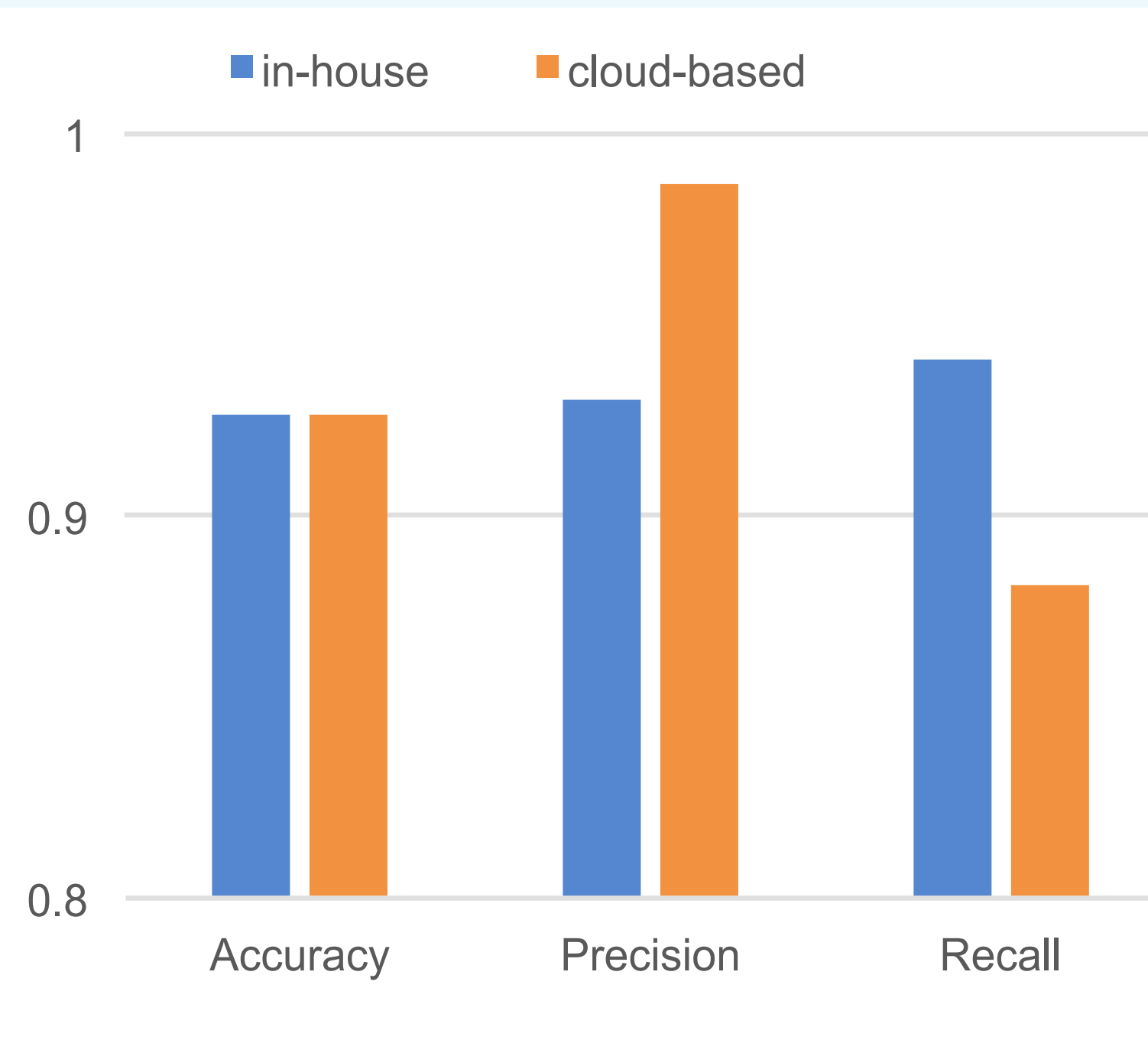
Classification Results from Machine Learning Techniques Tested

Building the Early Warning System

The ultimate goal of this research is to build an automated program that alerts violence intervention workers, community stakeholders, school personnel, and concerned residents about pending gang violence. When a gang's twitter activity crosses a threshold (whether level, types, or tone), the program will send a text message to those in the vicinity who can engage outreach, mediation, and other forms of de-escalation. By providing extra minutes, hours, or even days of warning, this system has the potential to save lives and reverse the current upswing in violence.

CONTACT: forreststuart@uchicago.edu & arriley@uchicago.edu