Protect
All
Children from
Trafficking™

Assessing Public Transit Rider Awareness of Sex Trafficking of Minors

John F. Betak[1], Ankit Sharma[1], and Christie L. Nelson[2]

[1]Collaborative Solutions, LLC, Albuquerque, NM
[2]Center for Discrete Mathematics and Theoretical Computer Science (DIMACS), Rutgers University

WeArePACT.org

PROJECT BACKGROUND

- Notice of Funding Opportunity (NOFO), March 2019
- Protect All Children from Trafficking™ (PACT, Formerly ECPAT-USA), May 2019
- PACT Notified January 2020
- Project Duration – July 2021 Through May 2025

OUR APPROACH

- Survivors-Informed Research-Based Social Media Campaign
- Six Transit Agencies
- Social Media Campaigns on Awareness
- Effectiveness of Survivor-Informed, Research-Based, Non-Traditional Anti-Trafficking Social Media Campaigns

PACT SURVIVORS' COUNCIL

- Lived-experience experts
- Inform PACT's programs, curricula, and public awareness efforts
- Consult and provide subject matter expertise to PACT staff during program development and evaluation
- Paid for their contributions in acknowledgment of the value that lived experience brings to our work

LITERATURE REVIEW

- 260+ Pieces of Literature in Our Search
- Categorized 245 Pieces of Literature
- 210+ Pieces of Literature Specifically Addressed Trafficking, Child Trafficking, & Trafficking on Transit
- 123 Discussed Child Abuse, Child Trafficking, Child Predators or Related Topics
- Limited Academic Literature on the Role of Transportation in Combatting Human Trafficking

TRANSIT AGAINST CHILD TRAFFICKING (TACT) CAMPAIGN

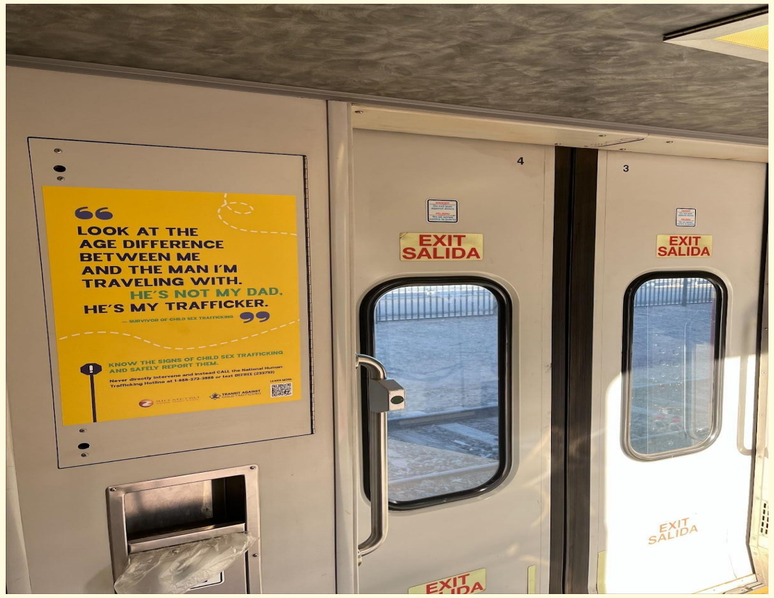
- Survivor-informed public awareness campaign deployed across six transit regions.

- Multi-channel rollout:
  - Transit interiors/exteriors, shelters, station posters
  - TV, radio, and digital media
  - Social media packages for local partners

- Agencies tailored visuals and messaging to local demographics and ridership patterns.

- Achieved strong visibility: 18 media placements and widespread station/bus/train coverage.

TRANSIT AGAINST CHILD TRAFFICKING (TACT) CAMPAIGN



RESEARCH METHODOLOGY

- **Social Media Scraping (EDA):**
  - **Collected 603,900 tweets using 260+ trafficking-related keywords/emojis.**
  - **Classified content using multiple ML models.**

- **Quantitative Awareness Assessment (Pre/Post):**
  - **Surveyed riders across six transit agencies.**
  - **Measured knowledge, perceived safety, indicator recognition, and reporting readiness.**

- **Predictive Modeling:**
  - **Built spatial-temporal models using publicly available datasets to forecast likely trafficking-related risk locations.**

Social Media Trends & Patterns

- Pre-pandemic dominant terms: #nowplaying, escort, sugaring terms.
- Post-pandemic dominant terms: #iheartawards, lolita, dateny/datela.

- Analysis of 7 major sporting events showed:
  - Top hashtags: #massage, #escort, #gfe
  - Suspicious tweets:
    - +48% year-over-year
    - +43% increase during events

- Combined keyword–location pairs revealed widespread use of #DateCity codes for escort promotion.

EXAMPLE KEYWORDS & HASHTAGS SCRAPED

Topic	Keyword, Hashtag	Definition	Primary Language (if known)
Child Sex Trafficking, Child Sexual Abuse	boylove	Indicates pedophile's gender preference	English
Sugaring	PapiDulce	sugar daddy in Spanish	Spanish
Escort, Prostitution, Sex related	transpinay	Filipino trans woman sex worker/entertainer	English
Sex related	Caldo de pollo	Spanish. Translates to soup chicken, which means having sex, or having sex with an older woman	Spanish

Tweet Classification Results

- Tested: Baseline, Gaussian NB, Bernoulli NB, Logistic Regression, Random Forest.

- Best-performing model: Random Forest
  - Accuracy: 0.966
  - Precision: 0.956
  - Recall: 0.912

- Demonstrates feasibility of automated detection of trafficking-related content at scale.

EXAMPLE EMOJIS SCRAPED

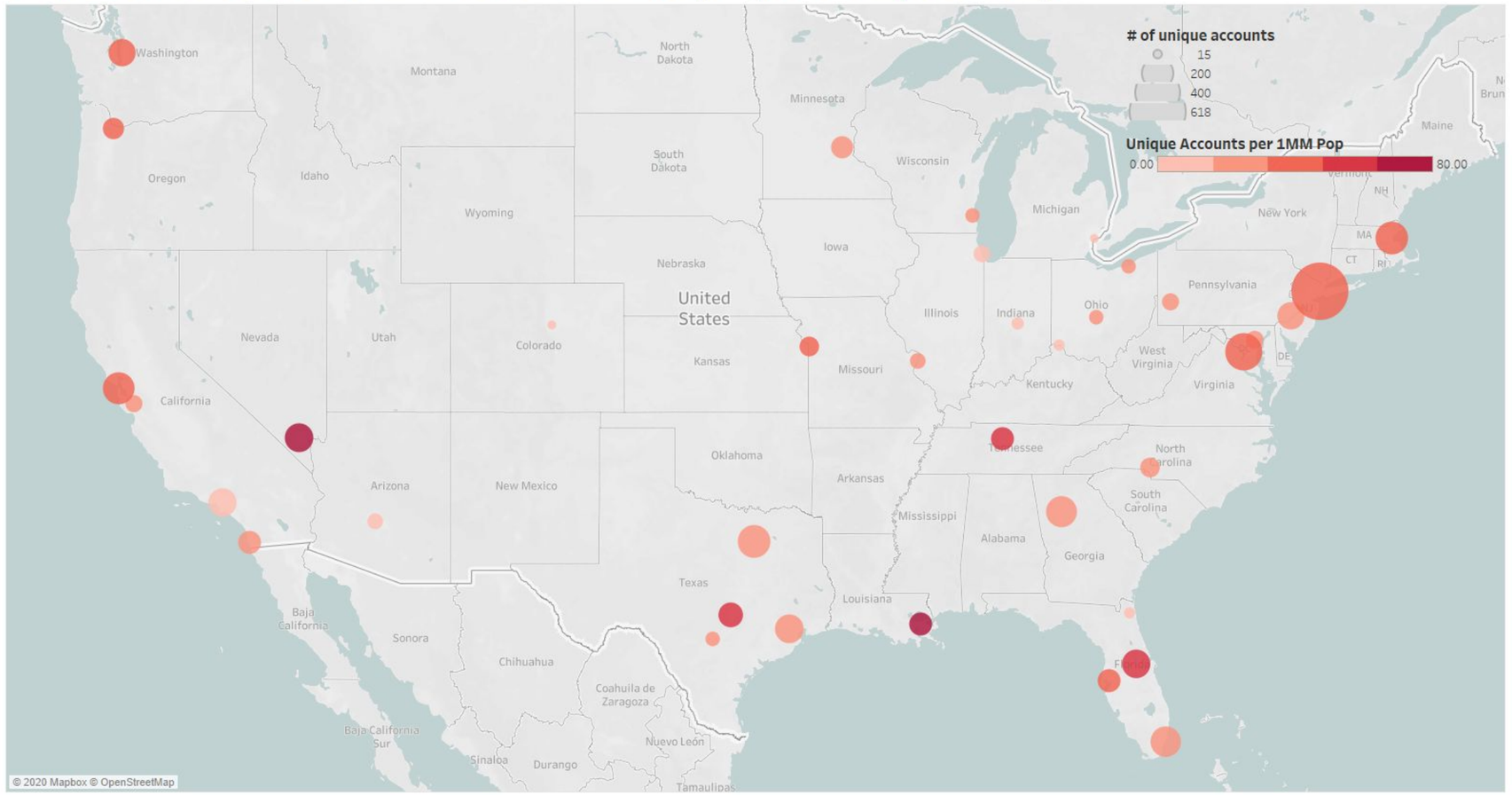
Topic	Emoji(s)	Definition	Type
Child Sex Trafficking, Child Sexual Abuse	💗	Indicates someone is a minor	Emoji
Grooming or Loneliness	🥀💔🔪	loss of loved one, depression or self-harm	Emoji
Child Sex Trafficking, Child Sexual Abuse	🌹📹💗✈️👑	translates to “Pay money to record having sex with an underage person who is being trafficked by a pimp”	Emoji
Child Sex Trafficking, Child Sexual Abuse	🔞	Under 18	Emoji

FOUR METHODS OF DATA TWEET CLASSIFICATION & RESULTS

Model	Accuracy	Precision	Recall	F Score
Baseline	0.743	1.00	0.743	0.852
Naïve Bayes - Gaussian	0.949	0.977	0.816	0.889
Naïve Bayes - Bernoulli	0.948	0.917	0.881	0.898
Logistic Regression	0.96	0.956	0.884	0.919
Random Forest Regression	0.966	0.956	0.912	0.934

FURTHER SOCIAL MEDIA DATA SCRAPING RESULTS

Unique Twitter accounts by for each city in the #date[city] searches (2018-2020)



PUBLICLY AVAILABLE DATA ANALYSES

- Integrated multiple publicly available datasets previously linked to trafficking-related activity:
  - Crime data
  - Socioeconomic indicators
  - Health indicators

  - Objective: Identify locations where multiple risk factors overlap, signaling potential hotspots or sourcing areas.

- Key statistical findings (national-level):
  - Poverty and trafficking strongly correlated ($p < 0.05$)
  - STI rates significantly correlated ($p < 0.05$)
  - Mental illness rates not statistically significant

Pre/Post Assessment Findings: Key Outcomes

- Pre/Post Awareness Assessment conducted across six U.S. transit agencies: CDTA (Albany, NY), DASH (Alexandria, VA), CapMetro (Austin, TX), CTA (Chicago, IL), TriMet (Portland, OR), Rio Metro (Albuquerque/Santa Fe, NM)
- Strong recall of TACT materials via TV, transit ads, and station signage.
- Portland & Alexandria showed improved indicator recognition.
- Many agencies reported declining perceived crime, suggesting increased feelings of safety.
- Digital channel visibility decreased across most systems.

Areas Requiring Improvement :Awareness Gaps Identified

- Recognition of vulnerable populations declined in several systems.
- Reduced perceived urgency of trafficking issues in the After-phase.
- Limited improvements in indicator recognition across most cities.
- Significant decline in knowledge of correct reporting actions in Chicago & Albuquerque/Santa Fe.
- After-phase respondents skewed older/higher-income, influencing overall results.

PREDICTIVE MODELING

- Modeled where & when child sex trafficking–related activity is most likely to occur near public transit.
- Integrated publicly available data: crime incidents, hotels/businesses, transit stops, events, neighborhood indicators.
- Tested multiple supervised + unsupervised frameworks to identify spatial and temporal hotspots.

Chicago (CTA): Data & Modeling

- **Data:**

8M+ crime records (2001–present) → filtered to 5K trafficking-related cases.
Census + business listings within 0.5 miles of Cicero Avenue.

- **Models tested: RF, SVM, XGBoost, Logistic Regression, Neural Network.**

- **Features: Lat, Long, Year, Month, Day, Hour.**

Chicago (CTA): Key Findings

- Neural Network highest accuracy (68.5%) & lowest error (MSE = 1.29).
- Clear route-specific clusters near:
  Pulaski & 46th Street
  Cicero & 72nd Street
- Peak risk windows:
  8–9 AM
  Early August
- Indicates consistent spatial-temporal patterns suitable for targeted interventions.

Chicago (CTA): Implications

- Data reveals stable, repeated hotspots along specific corridors.

- Can inform:
  - Operator alerts during peak hours
  - Focused patrols
  - Placement of awareness materials

- Future work: include additional businesses (massage parlors, entertainment hubs) to improve risk correlation.

Portland (TriMet): Data & Objectives

- Data: Crime records, hotels, TriMet stops, park-and-ride, massage parlor listings.
- 45,130 filtered trafficking-relevant offenses.

- Objectives:
  - Predict high-risk locations/times
  - Cluster neighborhoods
  - Link spatial factors with crime patterns
  - Provide actionable transit-focused insights

Portland (TriMet): Model & EDA Insights

- Random Forest classifier achieved 74.7% accuracy using:
  - Latitude, Longitude
  - Neighborhood
  - Month/Year

- Key insights:
  - Spatial dependency strongest predictor
  - Neighborhood context influences risk
  - Seasonal variations evident (month significant)

- Minority classes challenging due to data imbalance.

Portland (TriMet): Hotspots & Events

- K-Means clustering: 10 neighborhood clusters with similar spatial crime profiles.
- Crime concentrated within 0.5 km of 6 major TriMet transit centers (~5,200 cases).
- Event-driven spikes:
  - Festivals (Jazz Festival, Rose Festival, Blues Festival, Winter Light Festival, etc.)
  - Event periods show up to 100x higher daily offenses vs non-event days.

- Indicates short-term surges layered over long-term hotspots.

Cross-City Comparative Takeaways

- Both cities show clear spatial clustering near high-volume transit corridors.
- Chicago: Strong temporal regularity (peak hours + seasonal).
- Portland: Strong event-driven surges + persistent downtown hotspots.
- Predictive models demonstrate:
  - Public data can anticipate trafficking risk, not just describe it.
  - Transit agencies can adapt operator training, resource allocation, and awareness placement accordingly.

Integrated Conclusions:What the Study Demonstrates

- A survivor-informed awareness campaign, paired with rigorous statistical evaluation and predictive analytics, can improve public understanding of human trafficking on transit systems.
- The Awareness Assessment Tool captured rider perceptions, recognition of indicators, and readiness to act.
- Predictive models revealed consistent spatial and temporal risk structures that can guide campaign placement and operator preparedness.
- Together, these approaches provide a multi-dimensional view of trafficking risk—linking public perception with environmental and temporal patterns.

Key Implications for Transit Agencies

- Multimodal campaigns (digital, print, vehicle ads, stations) produce stronger retention than single-channel messaging.
- Localized tailoring to each city's demographics and ridership improves engagement and message relevance.

- Predictive spatial-temporal models can guide:
  - Strategic placement of awareness materials
  - Operator vigilance during high-risk hours/events
  - Focused outreach around persistent hotspots

- Integrating survey analytics + predictive modeling offers a replicable, data-driven framework for early identification, safer reporting, and improved prevention across transit environments.

ACKNOWLEDGEMENTS

- Rutgers MBS Externship and Practicum programs
- Transit Agencies
  - Capital District Transportation Authority - CDTA, Albany, NY
  - Rio Metro (Regional Transit), Albuquerque, NM
  - CapMetro, Austin, TX
  - DASH, Alexandria, VA
  - Chicago Transit Authority - CTA, Chicago, IL
  - Tri-County Metropolitan Transportation District of Oregon- TriMet , Portland, OR
- Fenton Communications
- RSG
- PACT
  - PACT Survivor Council members
  - PACT Summer Youth Interns
  - Lori Cohen
  - Yvonne Chen

THANK YOU

CONTACT US:

+718.935.9192 info@WeArePACT.org WeArePACT.org

Corresponding Author
John F. Betak,
john@collaborativesolutionsllc.com