

Project Title

Missed Detections: From Data to Actionable Estimates

Project Team Profile

- PI(s) Name(s), University:
 - Dennis Egan, Rutgers/CCICADA
 - Paul Kantor, Rutgers/CCICADA
 - Fred Roberts, Rutgers/CCICADA
- Project Start Date: July, 2017
- Anticipated End Date: June, 2020
- Project personnel:
 - Student (TBD)

Problem Statement

How can we estimate better the quantity of people, drugs, and contraband that are not detected crossing the border?

Can we discover the relative impact of various resources on the effectiveness of border operations?

Can we develop sets of data for these and related studies that can be published?

Beneficiary / End User Profile: Jobs

CBP/USBP Strategic Planning & Analysis
Directorate – Operations Officer

CBP/USBP Operational Requirements –
Assistant Chief

CBP Office of Information Technology – Software
Architect

DHS S&T Borders and Maritime Division –
Program Manager

Beneficiary / End User Profile: Desired Gains

- Better Allocation of Resources to Improve the Rate of Apprehending Illegal Migrants, and Confiscating Illegal Drugs and Contraband
- Better understanding of how to assign weights to various kinds of resources and measures of performance in relation to overall context and goals
- Provide access to key data (actual or synthetic) for researchers to use in developing and testing models for improving performance

Beneficiary / End User Profile: Pain Points

- CBP decision makers with operational responsibility will be able to understand and demonstrate success of investments in border security, detect changes and trends, and better assess risk
- CBP managers who develop requirements will understand how to weight different resources and performance measures to better achieve overall goals
- S&T managers will have a template for making data accessible for scientists to use in developing models and technologies

Products & Services

- TRENDFLAGGER Smart Spreadsheet Tool
 - Initiated in earlier work; will be extended for border data
- Principal Ray Finder to Estimate Flow Changes
 - Under development
- Methods to Synthesize Realistic Data
 - To be developed
- Extended Capture Recapture (ECR) Estimator
 - To be developed
- Data Envelopment Factor Analyzer
 - To be developed

Gains Created

- TRENDFLAGGER – Quickly detect positive and negative trends over time in easy to understand display (Metric: number of data sets for which it proves useful)
- Principal Ray Finder – Contextual estimates of percentage changes in the number of missed detections, without absolute scale (Metric: Adoption by CBP analysts)
- Methods to Synthesize Realistic Data – Enable access to data for scientific analysis (Metric: adoption by DHS S&T)
- Extended Capture/Recapture (ECR) Estimator – Get better estimates of missed detections (Metric: adoption by CBP)
- Data Envelopment Factor Analyzer – Allows optimally weighted comparisons of station inputs (resources) and outputs (performance) (Metric: adoption by CBP)

Pains Alleviated

- Missed early detection of positive and negative trends (Metric: Count the number of data sets for which TRENDFLAGGER is found useful)
- Poor/Uncertain estimates of Missed Detections (Metric: Principal Ray Analysis adopted by CBP)
- DHS Scientists unable to access key data sets (Metric: Methods for synthesizing realistic data adopted by DHS S&T)
- Poor/Uncertain estimates of Missed Detections (Metric: ECR Estimator adopted by CBP)
- Poor understanding of how different resources are best combined to improve operational performance (Metric: DEA Analyzer adopted by CBP)

Key accomplishments

- Kick-Off Meeting
 - Included representatives from key CBP and S&T stakeholder groups
 - Good consensus on work statement and transition path
- TRENDFLAGGER Smart Spreadsheet
 - Initiated in earlier work, refined on AIS data with application to USCG issues, planned extensions for border data
 - Prototype available for demonstration

Key accomplishments (2)

- Technical Analyses of ECR

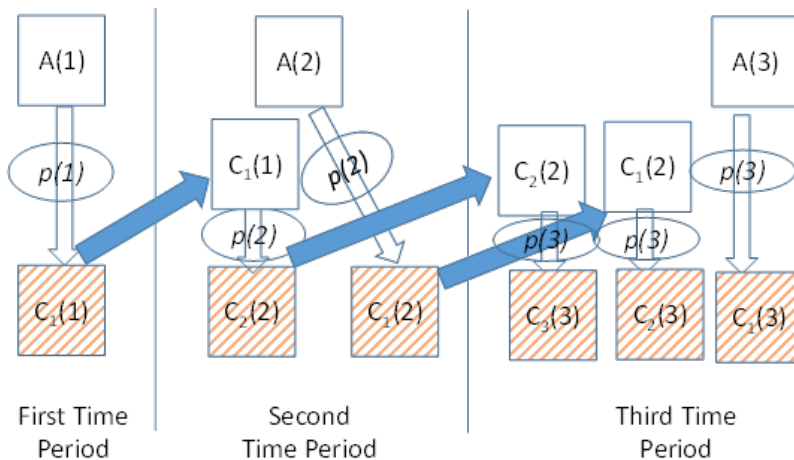


Figure 1. The Simplified Extended Capture-Recapture

- First Thoughts about Extensions
 - Multi-type Populations
 - Observable (e.g. nth-time Repeaters)
 - Hidden (e.g. Darwinian Models)

Transition Pathways

- Work with stakeholders (data stewards) to determine which data can be released and which will have to be anonymized or synthesized.
- Develop, refine, and test models on releasable or anonymized/synthesized data.
- Assist stakeholders to install project's software modeling and analysis tools in their environment, so stakeholders can test models on actual data.

Transition Engagement

- While we are early in this effort, here is what we can say:
- Key CBP stakeholders in the following roles were all represented at Kickoff meeting:
 - Data stewards
 - Data analysts
 - OIT software specialist
- All are eager to collaborate and see project succeed

Transition Challenges (if applicable)

- One key challenge is that data that are publicly releasable will only be useful for this research if complex relations among the anonymized data are carefully maintained
- This will make it challenging to develop tools that are useful enough to be disseminated to operational settings, providing essential “experimental feedback”

Conclusions

CBP is ready and engaged, so that the processes of defining synthetic data will begin promptly

There are several paths forward (ECR, Principal Ray Analysis, TrendFlagging) that can be pursued simultaneously to accelerate research

Disclaimer

- This material is based upon work supported by the U.S. Department of Homeland Security under Grant Award Number 2015-ST-061-BSH001. This grant is awarded to the Borders, Trade, and Immigration (BTI) Institute: A DHS Center of Excellence led by the University of Houston, and includes support for the project “Missed Detections: From Data to Actionable Estimates” awarded to Rutgers University, CCICADA. The views and conclusions contained in this document are those of the authors and should not be interpreted as necessarily representing the official policies, either expressed or implied, of the U.S. Department of Homeland Security.