

From RDS, to Capture-Recapture, to Multistage Probability Sampling: Data Gathering Methods for Human Trafficking Research

Sheldon X. Zhang, Ph.D.
Department of Sociology
San Diego State University
Email: szhang@mail.sdsu.edu



SAN DIEGO STATE
UNIVERSITY



turning knowledge into practice

High-risk Population



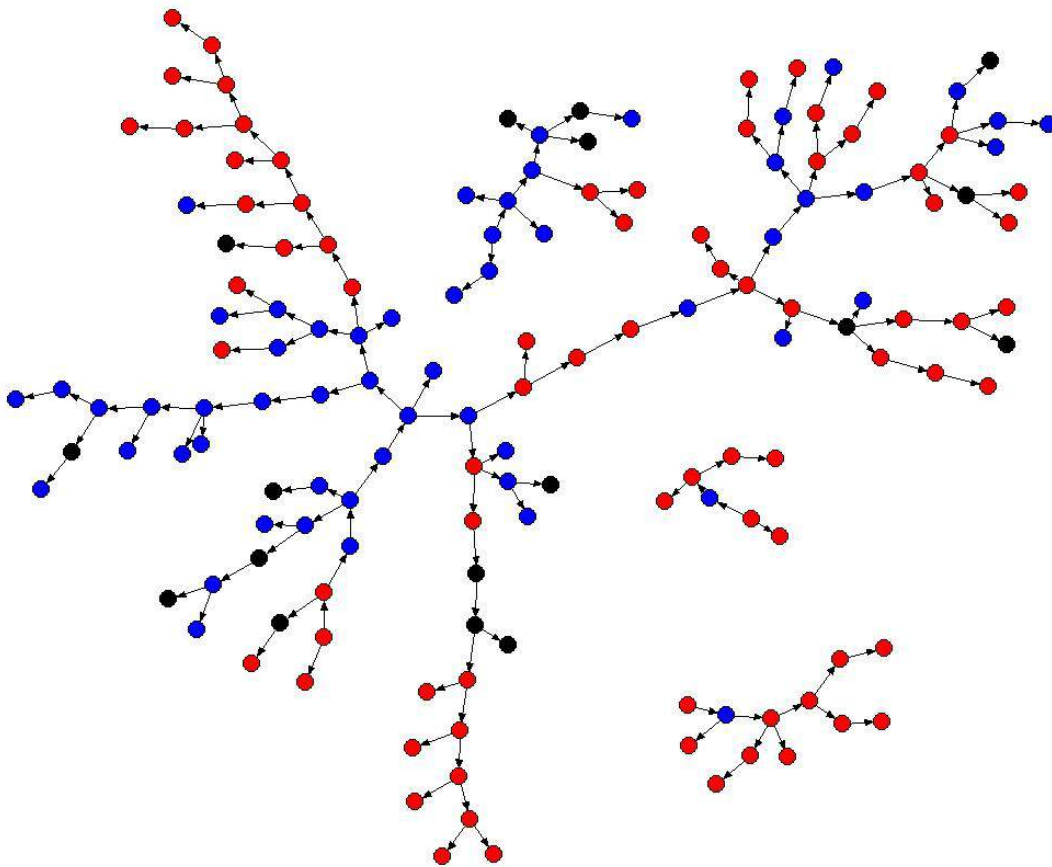
Populations Missed by Census



Unconventional Housing Arrangements

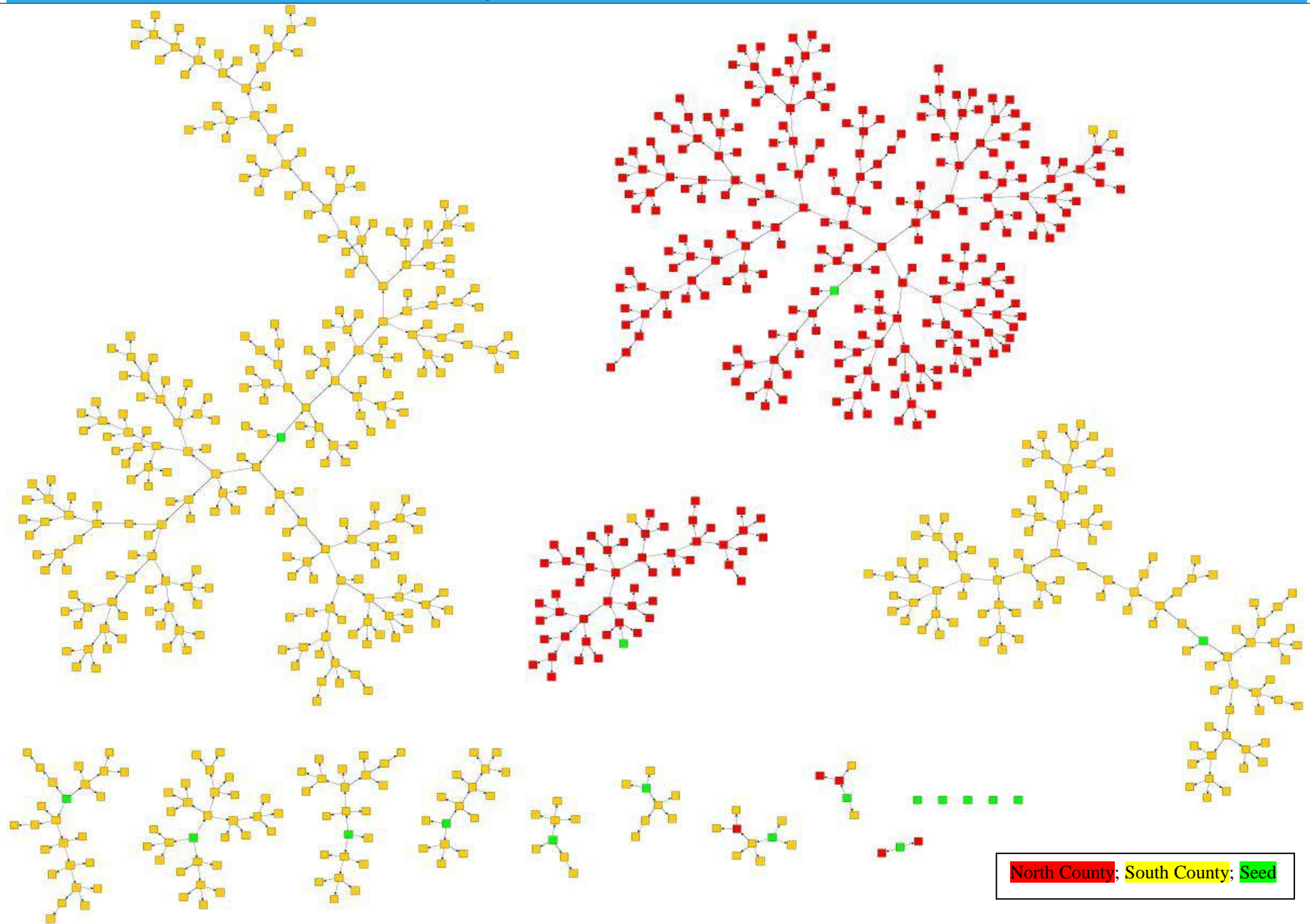


RDS Summary

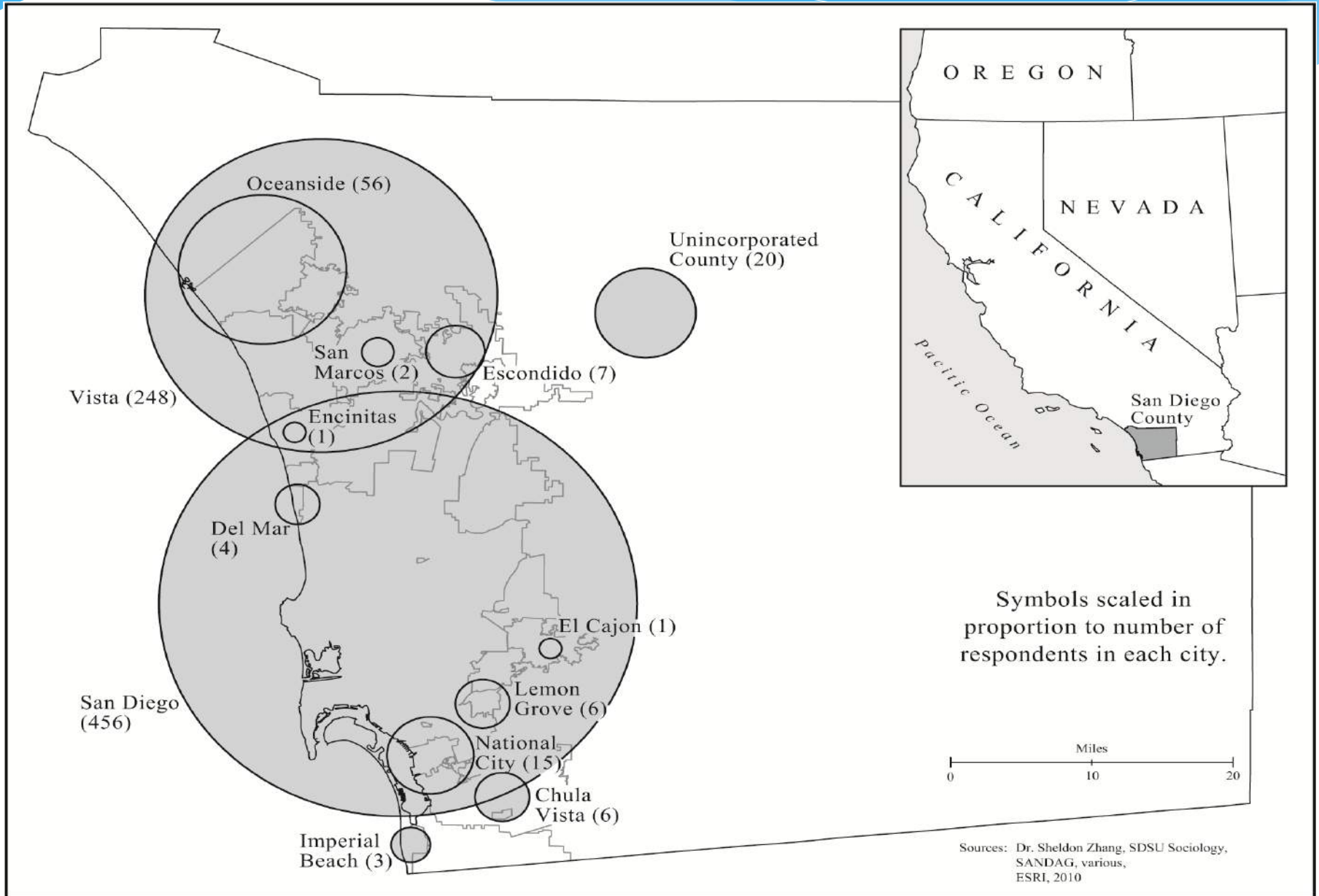


1. Starts with a small number of seeds
2. Seeds recruit new participants through social networks.
3. Who in turn recruit more participants...
4. Structured referrals should reach the edges of the migrant networks...
5. Multiple waves of recruits allow for unbiased estimation of the hidden population.

North and South County Divide

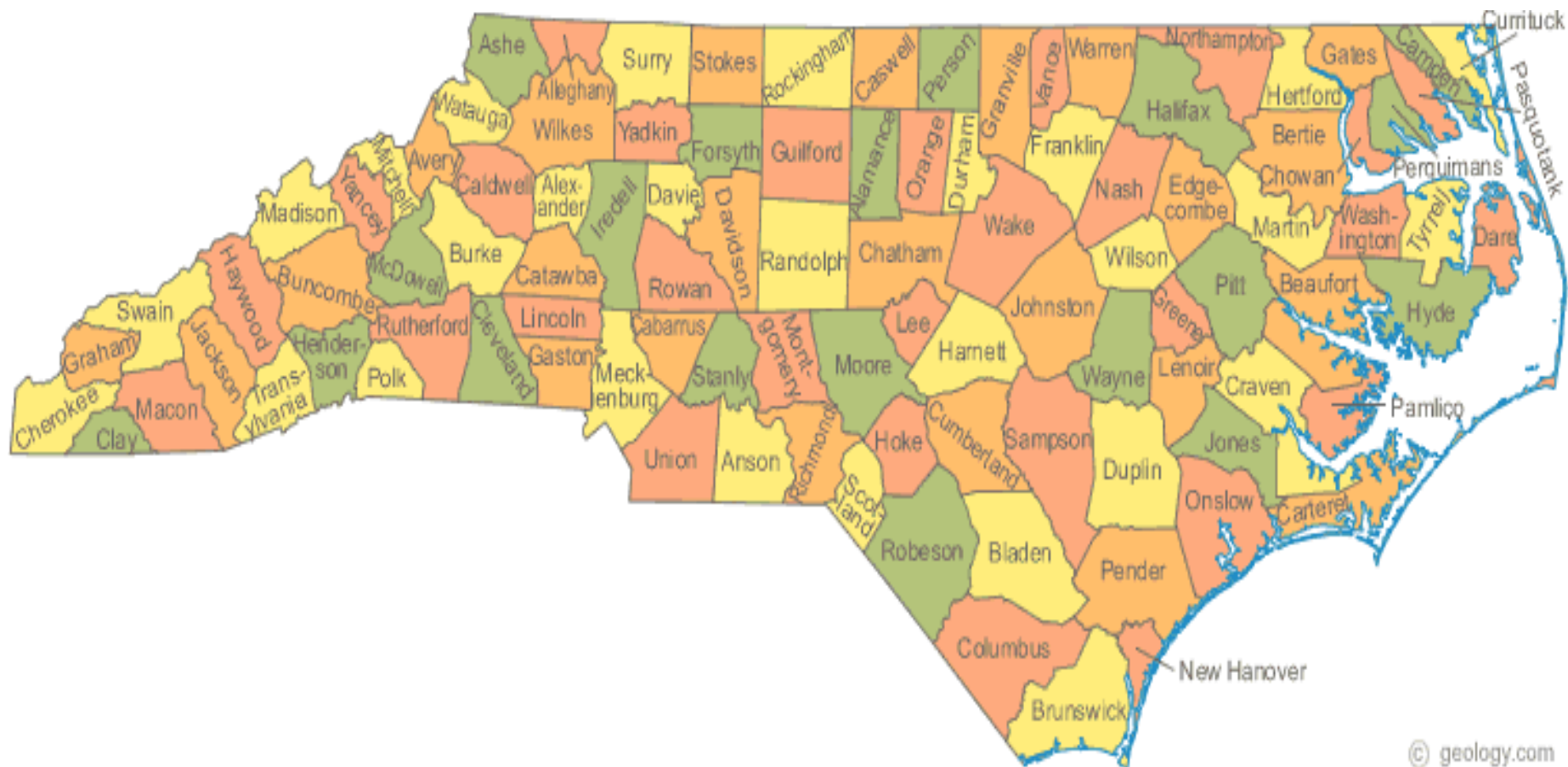


San Diego Map and RDS Fielding Activities



Current Study—NIJ Funded

- * Estimate the prevalence of labor trafficking (gross exploitation) victimization
- * Labor market specific (target population → migrant farmworkers)
- * North Carolina (statewide)
- * Multistage sampling
- * **Main challenge:** probability-based sampling requires a sampling frame



First Stage: Primary Sampling Units (PSUs)	
Sampling Frame	Area segments in NC
Stratification	None
Type of Selection	Probabilities Proportional to Size Measure ¹
Sample Size	20 areas
Second Stage: Segments (groups of census blocks)	
Sampling Frame	Segments (groups of census blocks)
Stratification	None
Type of Selection	Probabilities Proportional to Size Measure ¹
Sample Size	192 (9 to 10 segments per selected PSU)
Third Stage: Dwelling Units (DUs)	
Sampling Frame	Field Enumerated DUs using GeoFrame methodology ²
	Expected 42,240 Field Enumerated Dwelling Units (stratified during enumeration)
Stratification	1) High (expected distribution: 1%) 2) Medium (expected distribution: 6%) 3) Low (expected distribution: 93%)
Type of Selection	Systematic Random Sample (oversample High and Medium strata)
Sample Size	Selected DUs: 3,122
	Screened: 2,529 (81% screening rate)
	Eligible: 804 (32% eligibility rate) ³
	Complete: 724 (90% completion rate) ⁴

Strategies to Enumerate

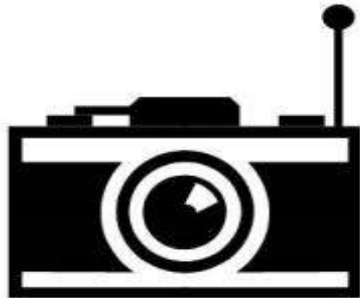
- * Conventional Census enumeration strategies don't work.
- * We will use GeoFrame™, a method of field enumeration developed at RTI:
 - * Captures digital images with GPS coordinates.
 - * Use digital images and GPS coordinate serve as a surrogate to what would be a household address or description with traditional field enumeration.

GeoFrame in a nutshell

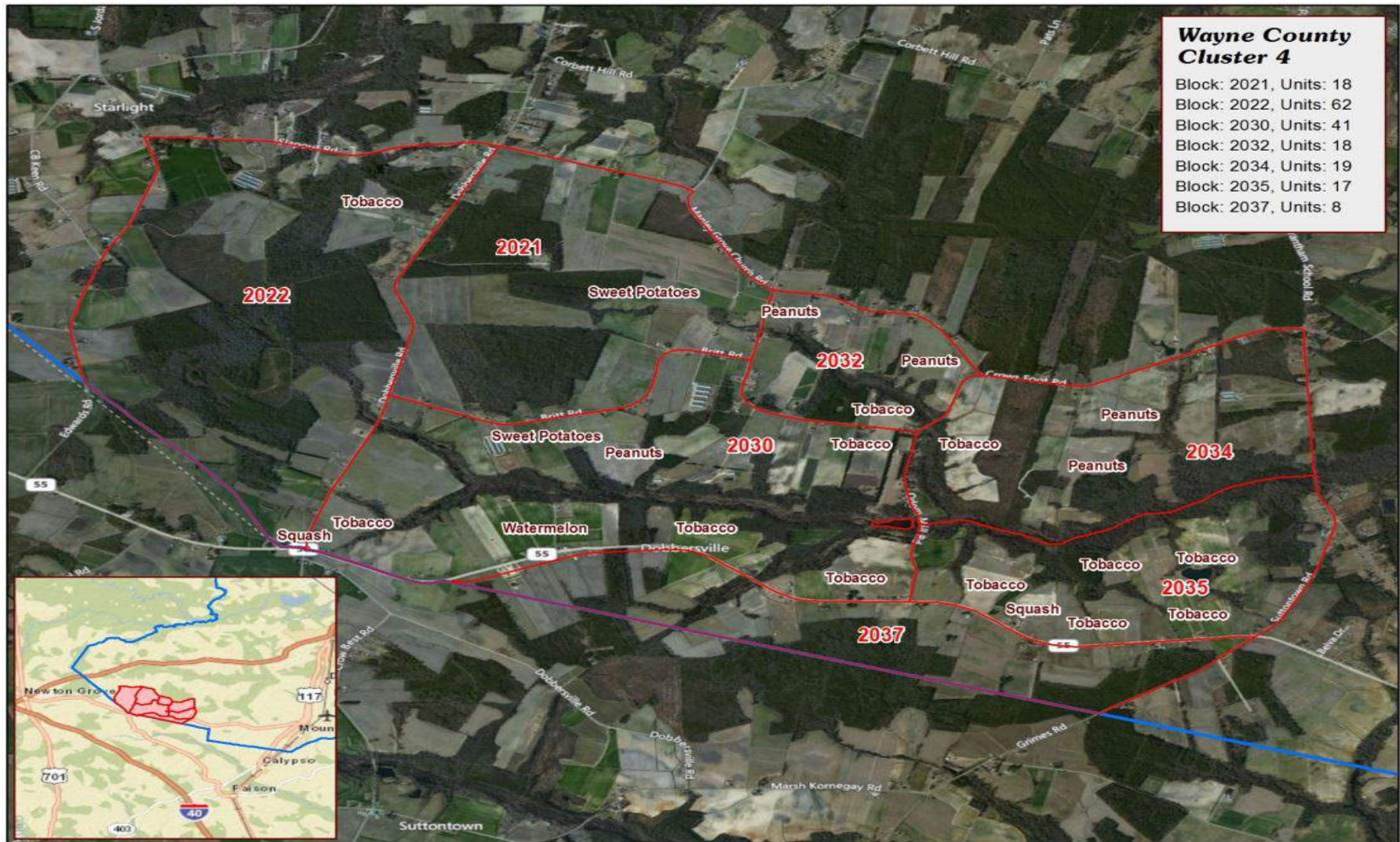
Geo-Positioning Satellite



Lat: $XX^{\circ} XX' XX''$
Lon: $YY^{\circ} YY' YY''$



Wayne County Cluster 4, the target block (#2030) and all of the blocks that touched it





Lister Path of Travel:
Location, ID & azimuth:
Photo Field of Vision:



Adaptive Sampling: The Next Frontier!

population graph

