

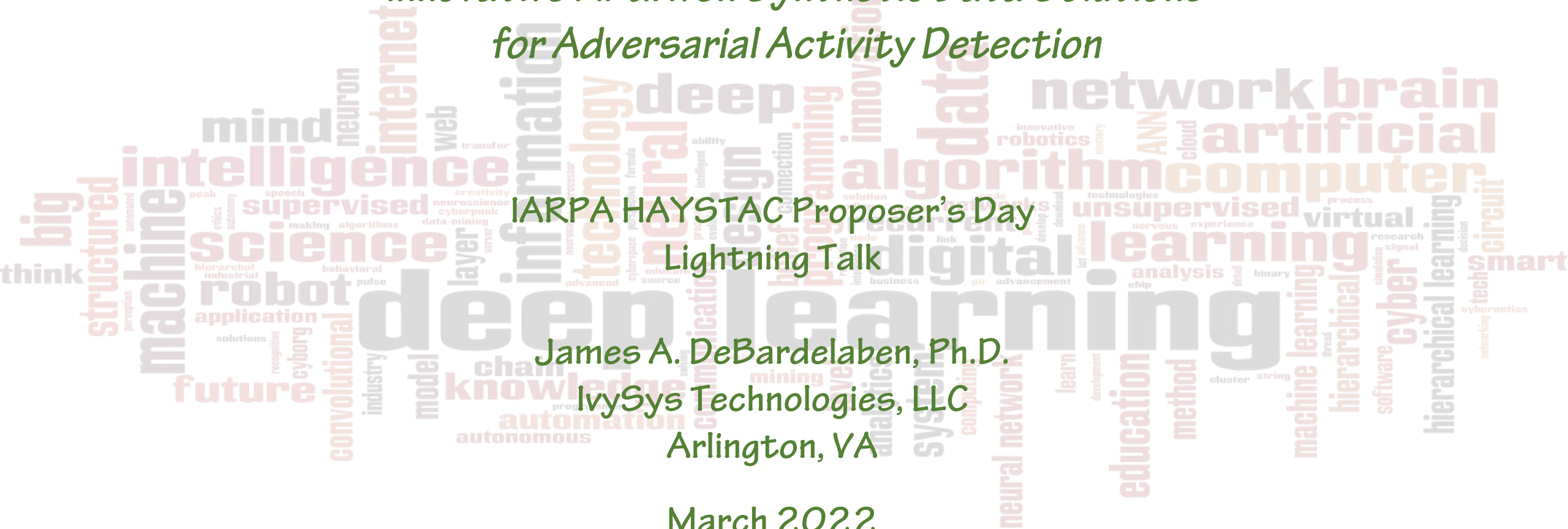


*Innovative AI-driven Synthetic Data Solutions
for Adversarial Activity Detection*

IARPA HAYSTAC Proposer's Day
Lightning Talk

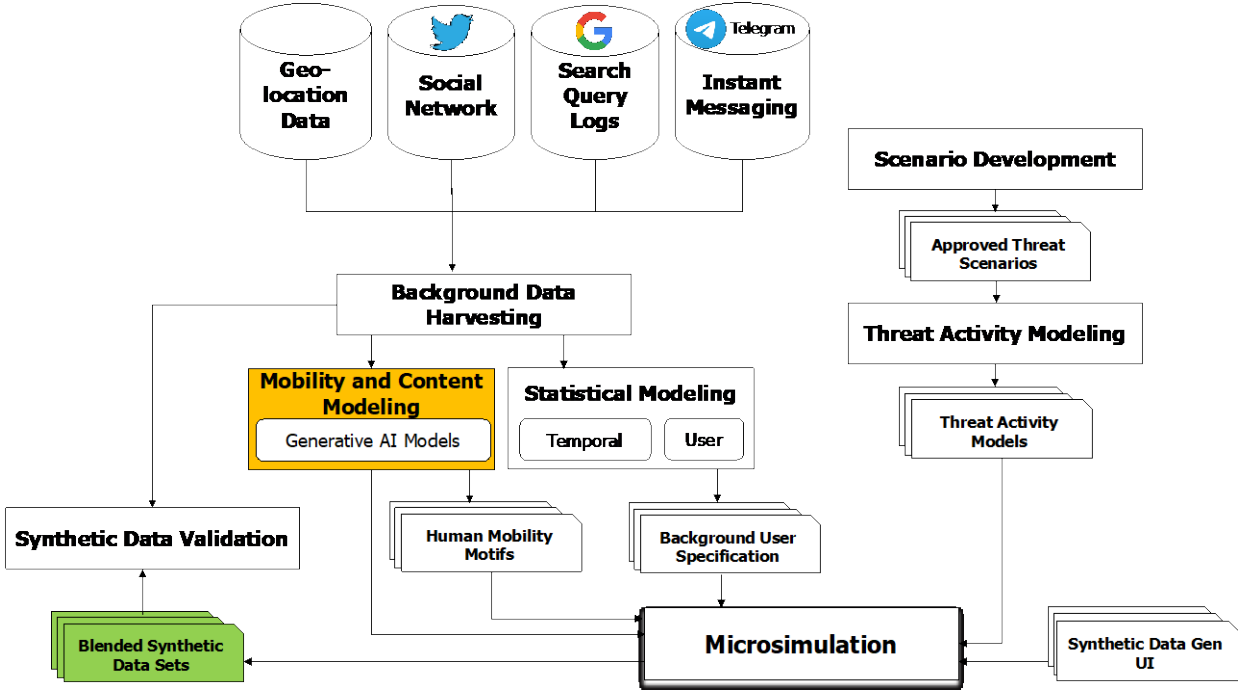
James A. DeBardelaben, Ph.D.
IvySys Technologies, LLC
Arlington, VA

March 2022



Overview

IvySys provides innovative, high-quality synthetic data solutions and services that enable organizations to rapidly train AI models to discover threats hidden in big data to keep our nation safe.



- Example Synthetic Data Sources
- ✓ Geolocation Data
 - ✓ Social Media Posts
 - ✓ Instant Messaging
 - ✓ Group Chats
 - ✓ Web Search Logs
 - ✓ Online Forums
 - ✓ Blogs
 - ✓ Emails

AWS cloud-native application leveraging large-scale, generative deep learning models

Corporate Snapshot



- *Founded by Dr. James A. DeBardelaben*
- *Headquartered in Arlington, VA*
- *Top Secret Facility Clearance*
 - *75% of Technical Staff are TS/SCI Cleared*
- *Primary NAICS: 541715*
 - *Research and Development in the Physical, Engineering, and Life Sciences (Except Biotechnology)*
 - *Secondary NAICS: 541330, 541511, 541512, 541519, 541690, 541990*
- *Certified Small Disadvantaged Business (SDB)*
- *CAGE: 4CS09*
- *DUNS: 608730714*



Trajectory Generation Problem

Generative Mobility Model M

- Generate n synthetic trajectories $T_M = \{T_{a_1}, \dots, T_{a_n}\}$,
where T_M specifies movements during a certain time period,
for very large number of n independent agents a_1, \dots, a_n
- *Time-ordered sequence* $T_{a_i} = \{p_1, p_2, \dots, p_k\}$
Spatial-temporal points $p_i = (t, x, y)$ describe the k locations visited by agent a_i

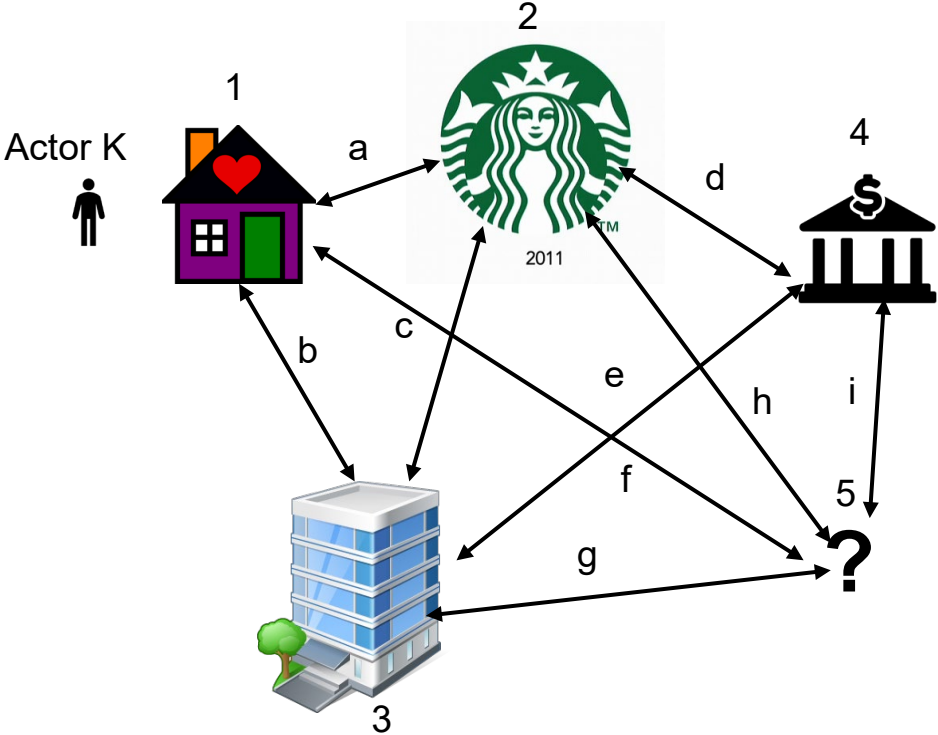
Leverage DARPA MAA Program

- **Synthetic Transaction Data Generator (STDG) for DARPA Modeling Adversarial Activity Program**
 - Generated unclassified temporal transaction graphs containing WMD threat and Background activities
 - Generated data sets with 1 million entities (nodes) and 1 billion transactions (edges)
 - Computed Signal-to-Noise ratio (SNR) metrics in real-time on generated data and make user configurable
 - Matched background data statistics to real-world data sets (e.g., geolocation, communications, financial, logistical)
 - Developed STDG in modules:
 - **Synthetic Network Generator (SNG),**
 - **Synthetic Transaction Graph Generator (STGG)**
 - **Data Blender**
- **Relation to DARPA MAA Vision**
 - The STDG provided synthetic transaction data to drive the development of mathematical techniques that integrate and analyze heterogeneous data sources to enable high confidence indications & warnings of WMD activities
 - Avoided the privacy and classification issues that can be associated with real-world data.

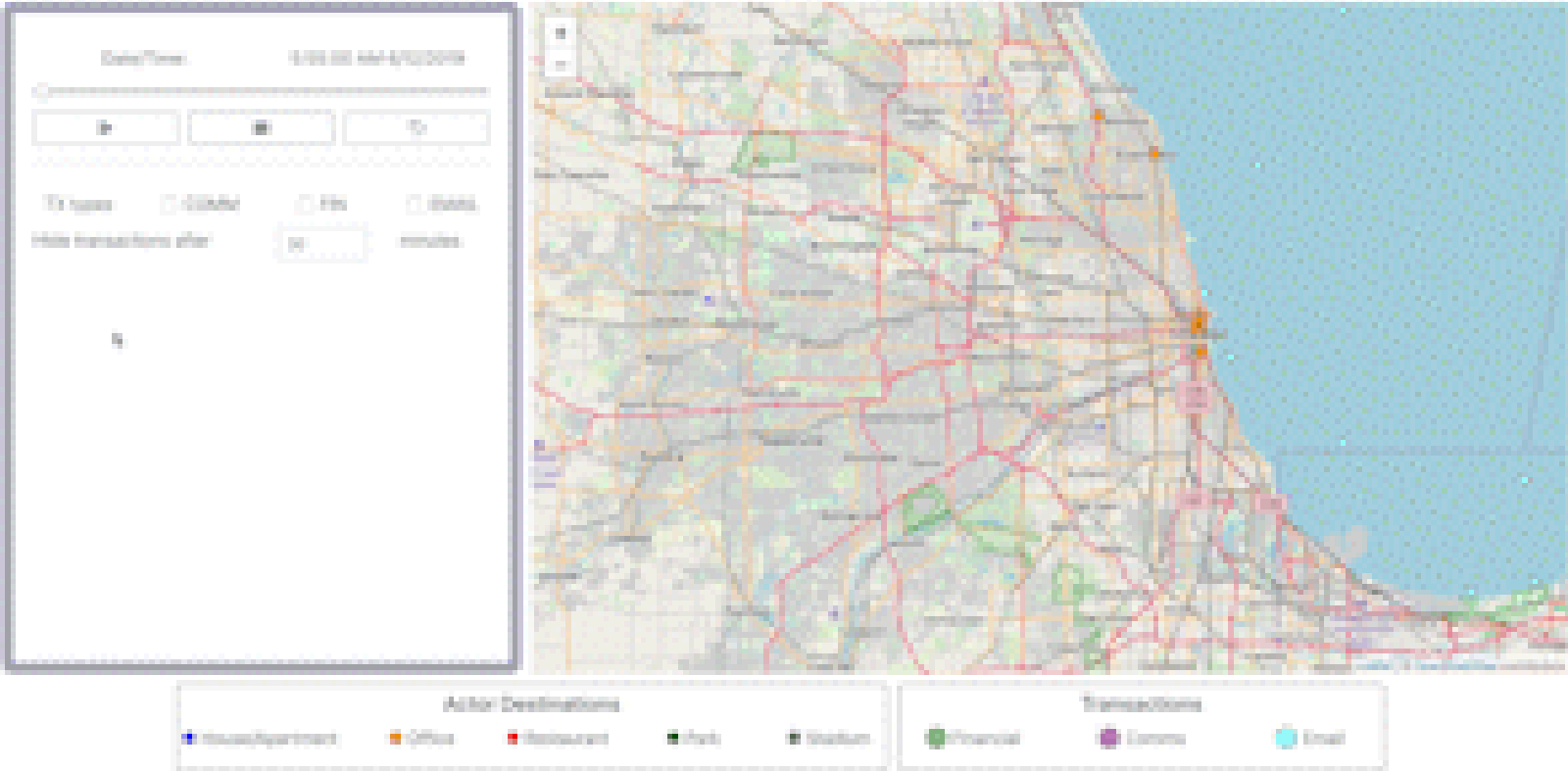
STDG Human Movement Simulator

- *Actors in the movement simulator were provided locations for home, work, frequent shopping, family members, etc.*
- *Plan to expand with more detailed, patterns of life, mostly city-level movements.*
- *Individuals travel between cities via car, bus, train, or plane.*
- *Most individuals (agents) at their place of work during typical work periods of the work location.*
- *Model venues that serve as less frequent but highly attractive places to congregate, like sports stadiums or shopping malls.*
- *Given the actor's location at any time, the simulator will be equipped to attach the account's holder's location to whichever transactions the user deems appropriate.*
- *Generate slight deviations of agent locations to represent roaming within a work campus or walking to nearby lunch locations.*
- *Location can dictate which accounts (e.g., financial, social media, etc.) are used for transactions at a given time.*

Human Movement Model



Human Movement Model Demo





Contact Information

James A. DeBardelaben, Ph.D.

President and CEO

IvySys Technologies, LLC

4501 N. Fairfax Drive, Suite 920

Arlington, VA 22203

Phone: 703-414-5665 x111

Email: jd@ivysys.com