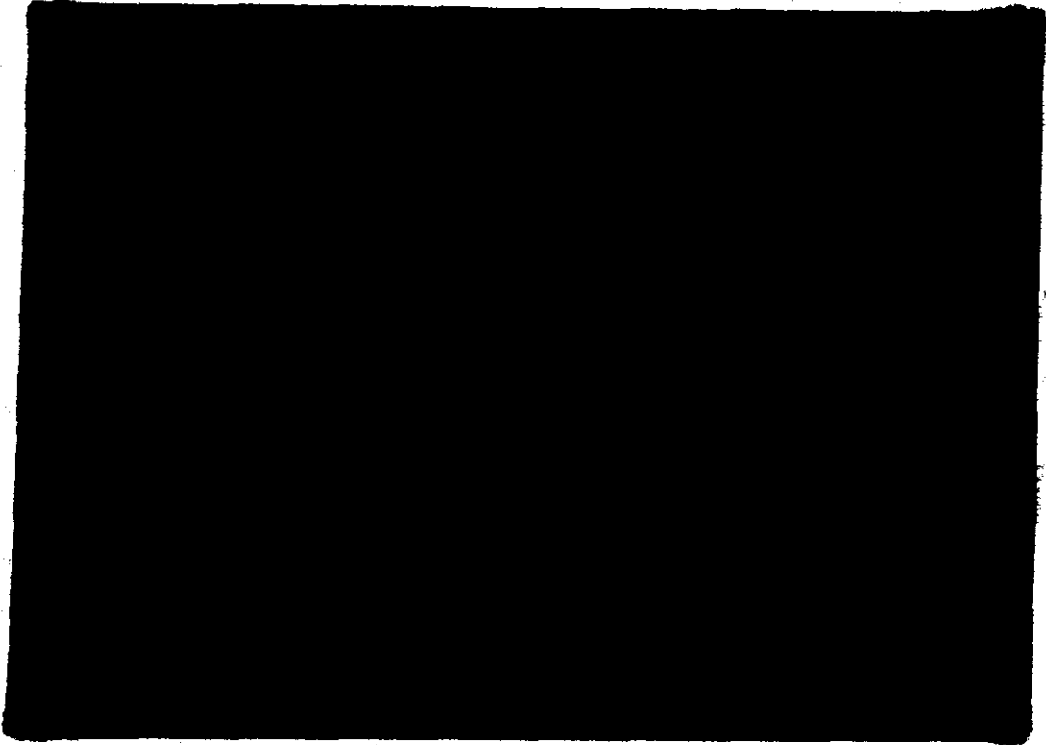


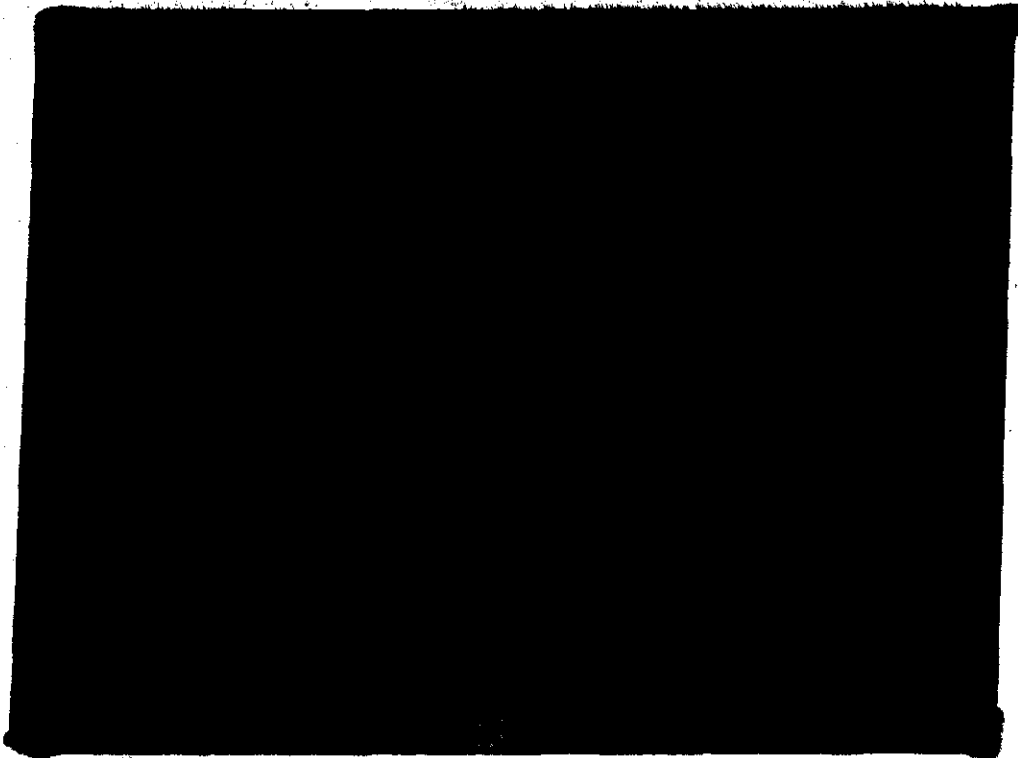
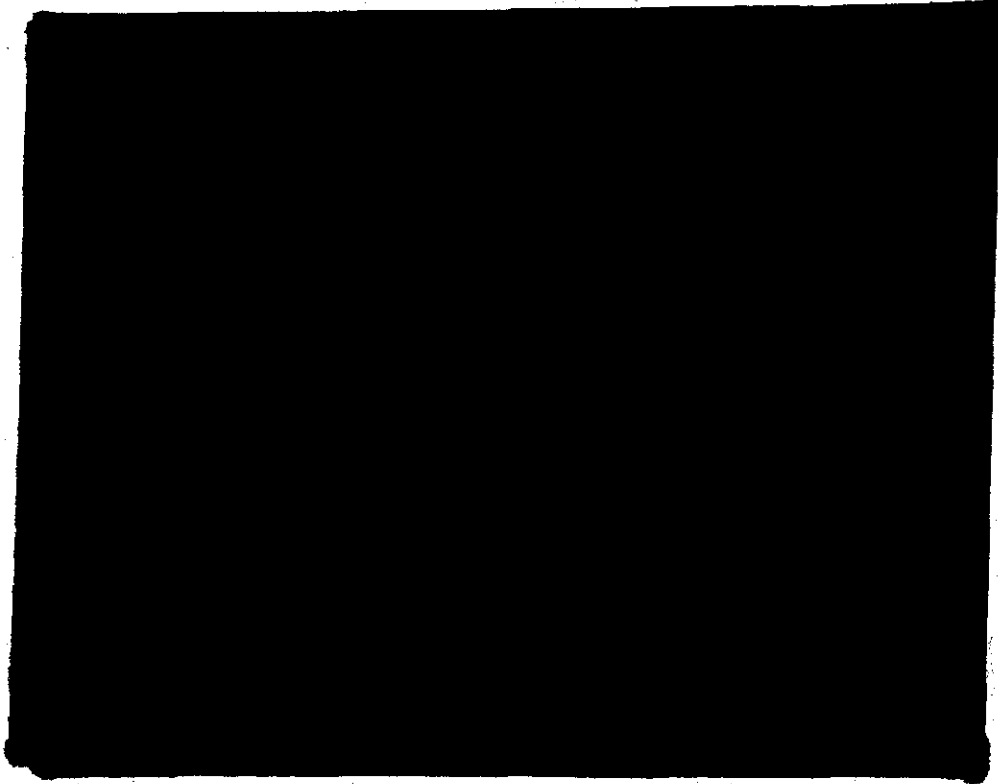


# Passenger Threat Assessment Break-out Session

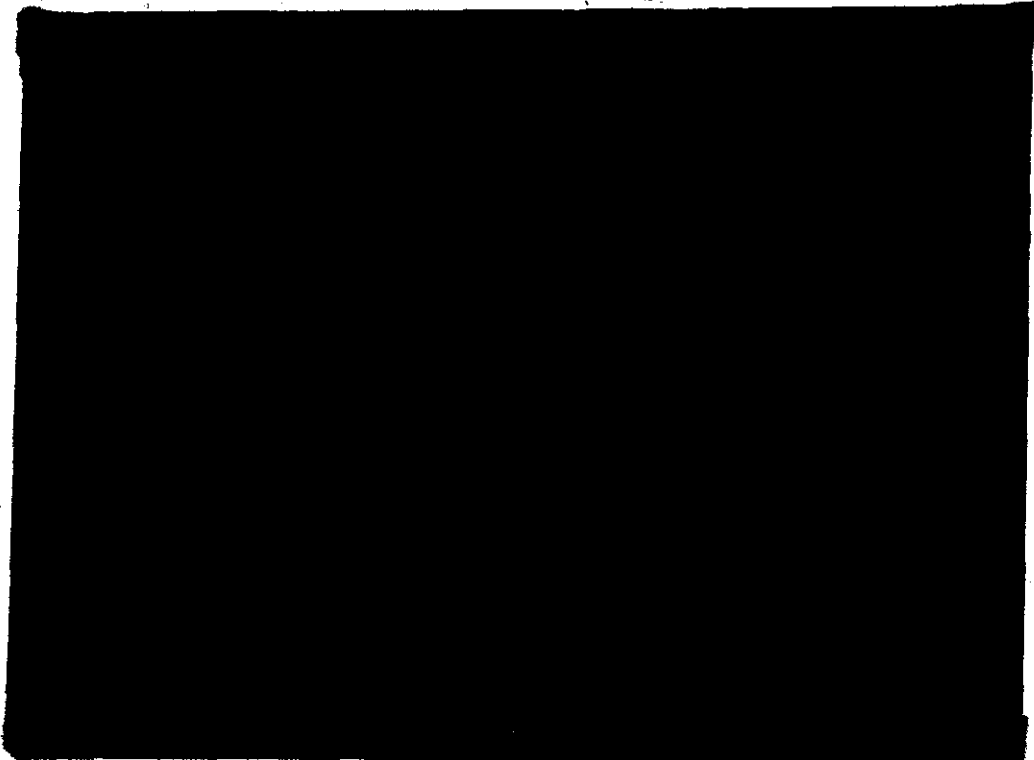
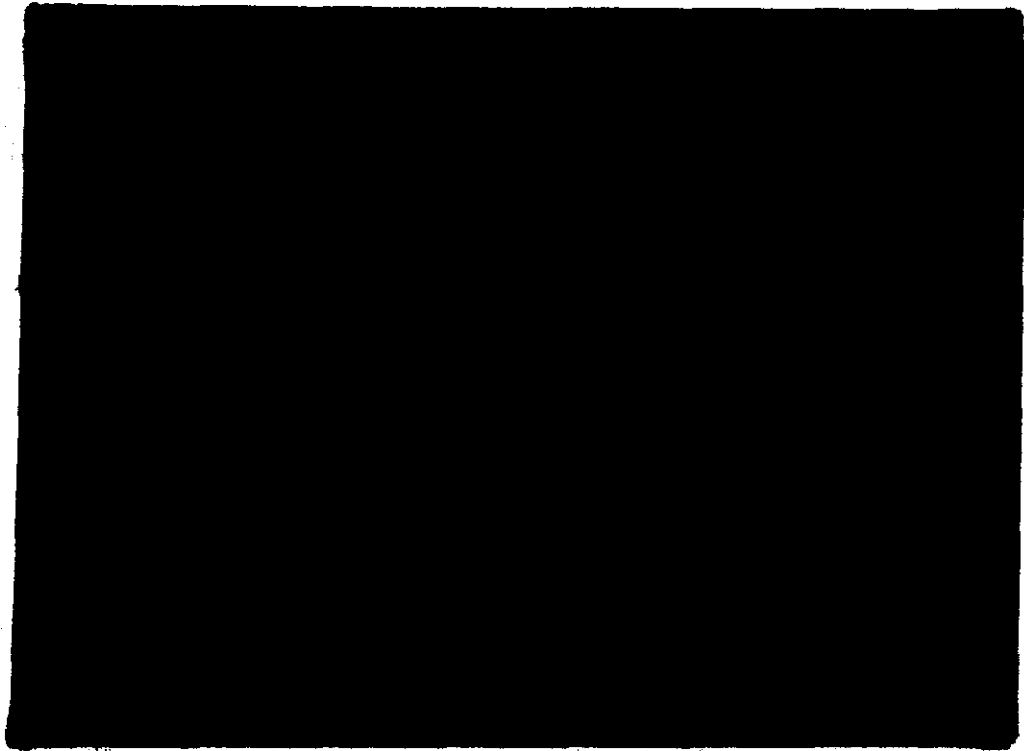
Session Co-Leads  
Paul Ruwaldt, FAA Tech Center  
Tom Hinke, NASA Ames



2/3/2004



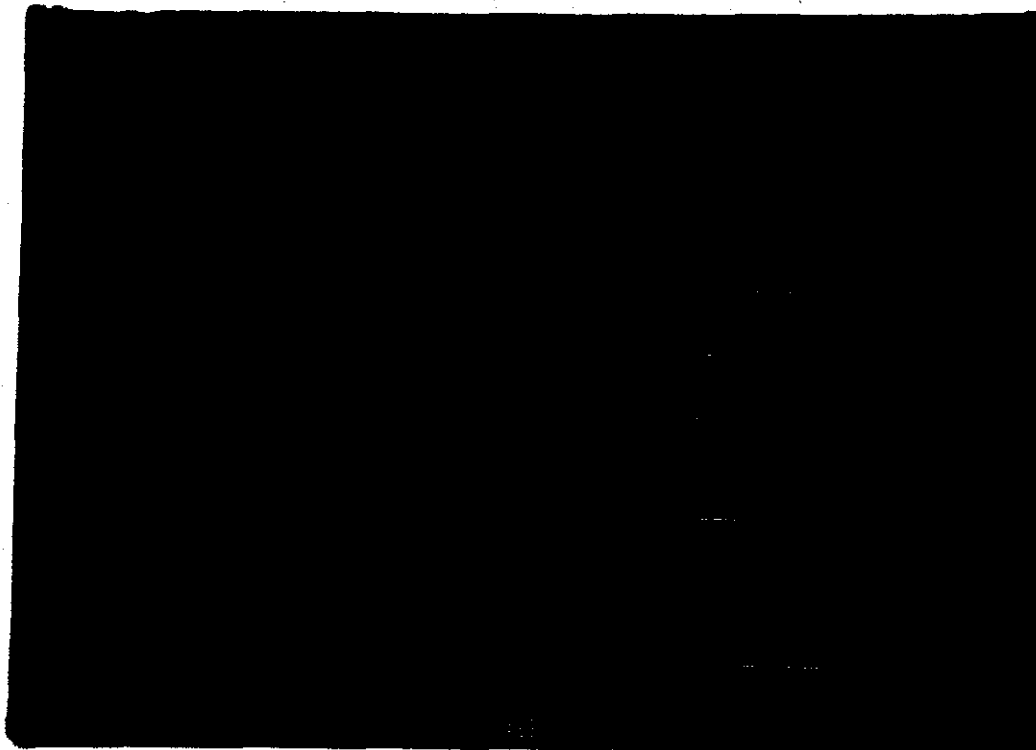
2/3/2004

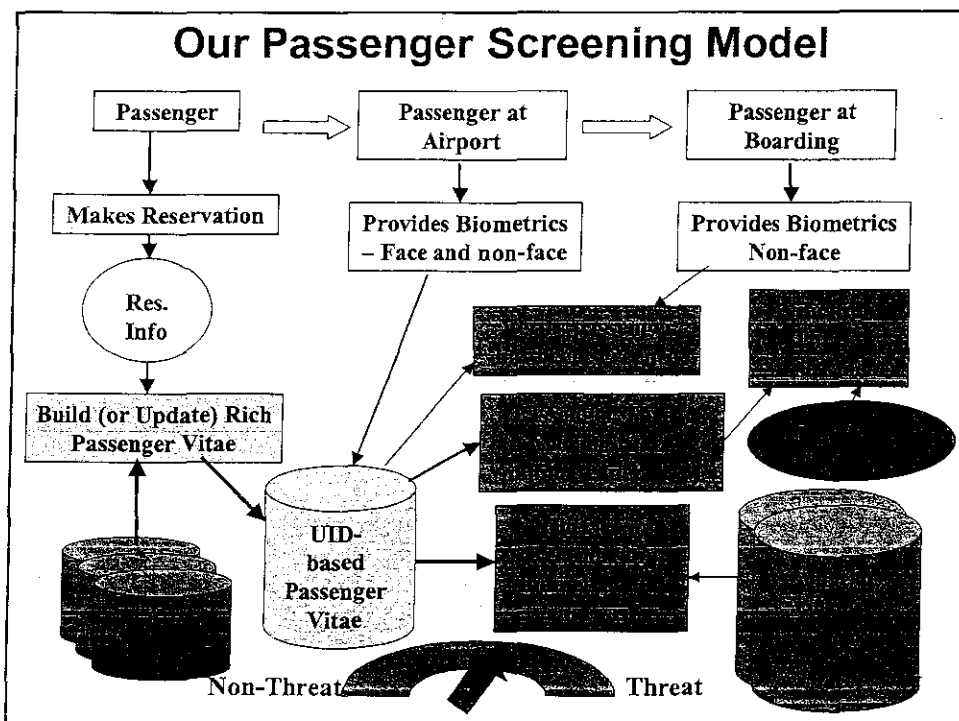
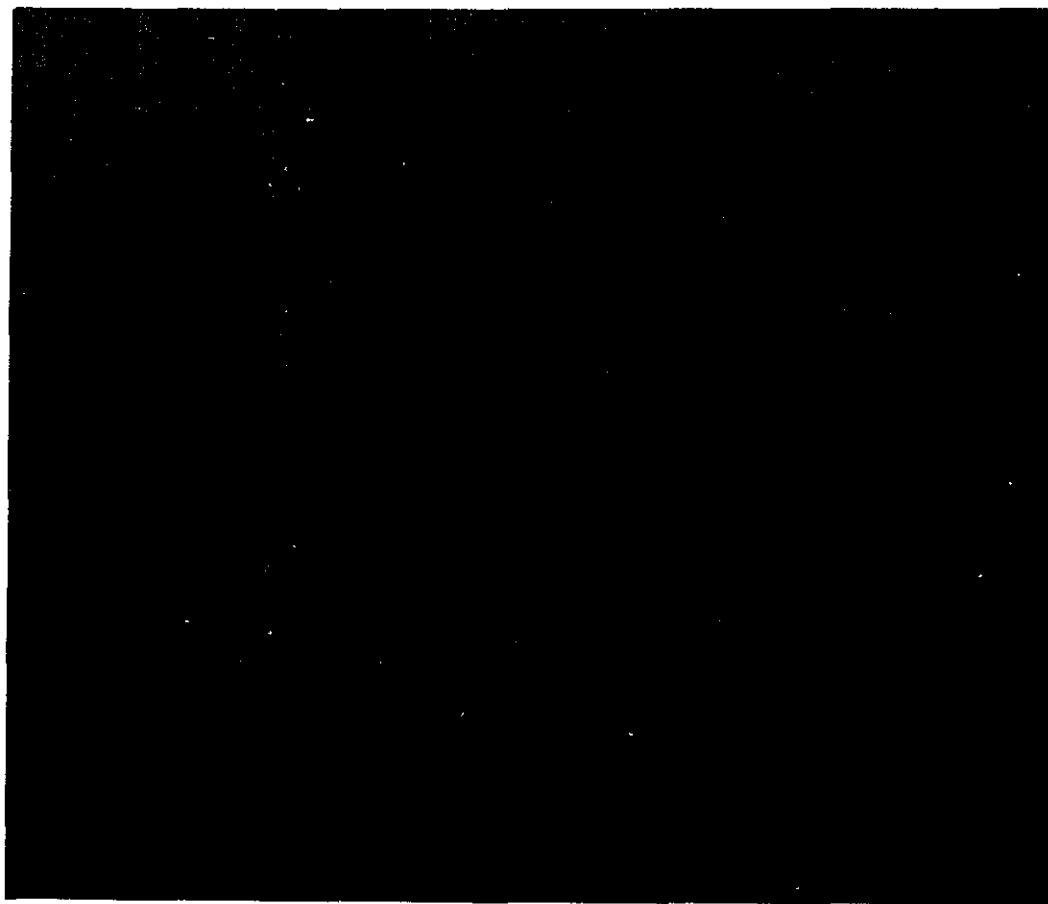




## Review of NASA Near-Term Initiatives in Passenger Threat Assessment

Tom Hinke, NASA Ames Research Center







## NASA Near-Term Initiatives in Passenger Threat Assessment

- **Aviation Security Lab:** Established to explore the application of information technology to passenger threat assessment for secure passenger flow
- **Knowledge Acquisition Infrastructure:** Developed to provide a foundation for researching the application of information technology to passenger screening and threat assessment.
- **Data Analysis/Mining:** Performing research into the application of data mining and machine-learning technologies in support of automated passenger threat assessment.
- **Security Reporting:** Performing analysis of security reports current being received by Aviation Safety Reporting System(ASRS) and encouraging use of ASRS for future security reporting.

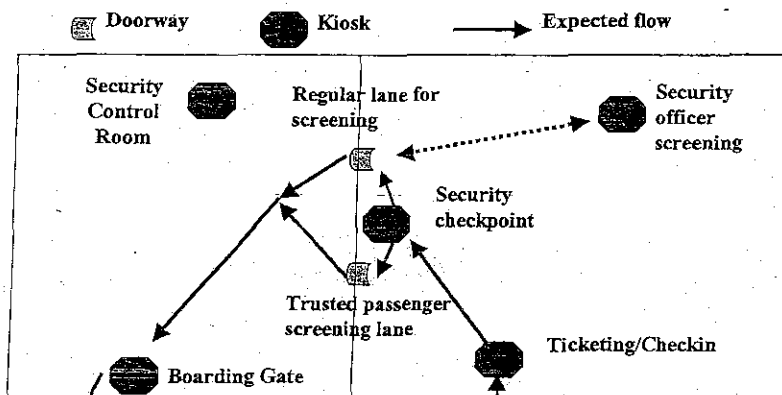


2/3/2004



## Established Aviation Security Lab

- Allows testing of various mixes of information technology in passenger flow



2/3/2004





## Implemented Knowledge Acquisition Infrastructure

- Based on user-defined world model (a semantic model)
- Various data extractors access data
  - Reservation information
  - Government agencies – watch lists
  - Commercial sources – Lexis/Nexis
- Provides rich integrated data environment about
  - passengers
  - suspected terrorists
- Initially using commercial Voquette software



2/3/2004





## Data Mining Research

- Data mining seeks to automatically discover hidden relationships in large databases
- Initiated work to apply various data analysis/mining approaches to threat assessment problem
  - Expert Systems – rule based
  - Inductive Learning – train on normal and threat passengers
  - Anomaly detection – mine for deviations from norm
  - Link analysis – look for connections
  - Clustering – find clusters of safe passengers such as business travelers or tourists
  - Hybrid mining – combine two or more approaches.
  - Group threats – explore methods for detecting groups of high-threat passengers on the same flight, or across multiple flights



2/3/2004



## Expert Systems Approach

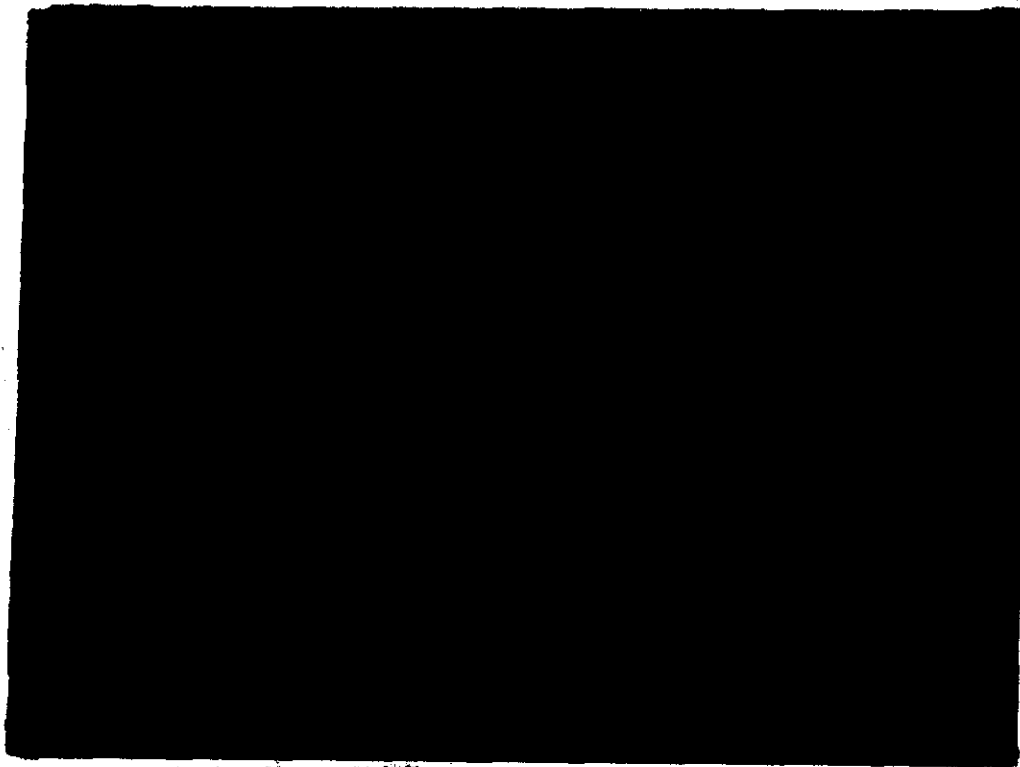
- Human experts on aviation security are asked how they would assess the threat level of a passenger
- Their methods for threat assessment are encoded into a set of rules
- For example, the experts might say that a passenger who pays cash for a one-way first-class ticket is more likely to be a threat.
- This approach was used to build CAPPS



2/3/2004







## Anomaly Detection Approach

- Uses data about “good” passengers only
- Builds a model of what passengers typically looks like, and flags “anomalous” passengers who don’t fit the model
- For example, a passenger who paid cash for a first-class one-way ticket would probably be anomalous
- Selected passengers would be subjected to additional screening
- We are initially using GritBot from Ross Quinlan’s Rulequest Research





## Link Analysis Approach

- Two people are linked if (for example)
  - They have the same previous address (from LexisNexis data)
  - They flew together (from NWA data)
  - One made a phone call or sent an e-mail to the other (don't have this data yet)
- A passenger is more likely to be a threat if he/she is linked to a suspected terrorist, and less likely to be a threat if he/she is linked to someone who is believed to be safe
- Threat scores are propagated across links based on the weights of the links



2/3/2004



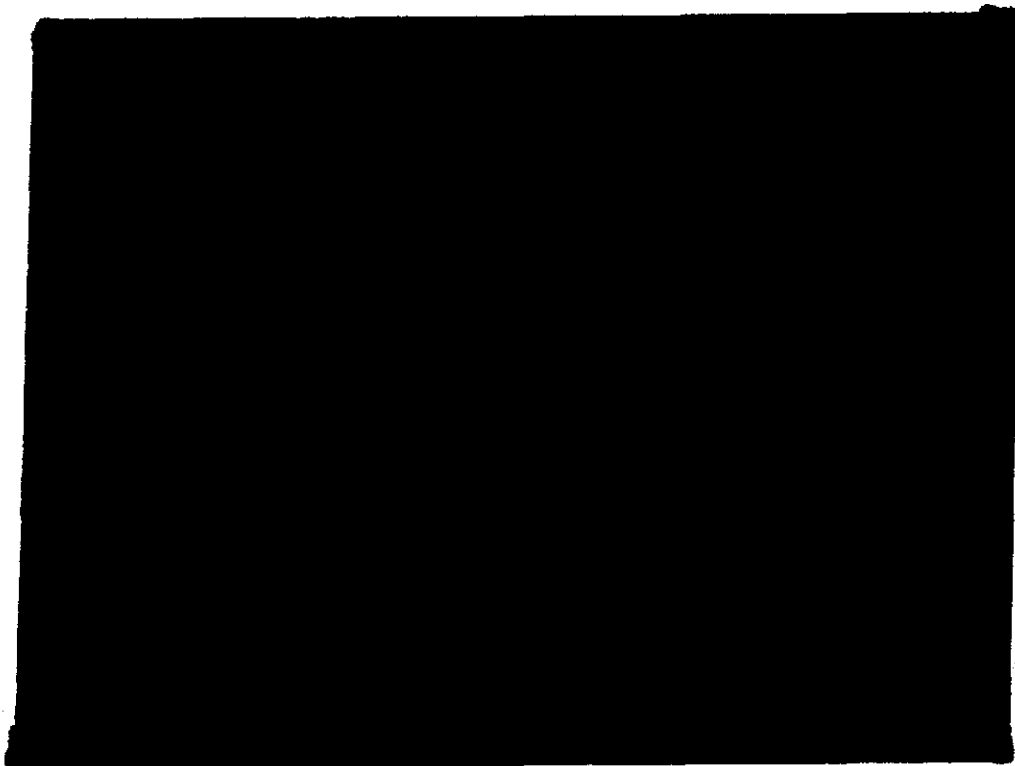
## Clustering Approach

- Use a clustering (or unsupervised classification) algorithm to find clusters of similar passengers
- Objective is to discover clusters that correspond to classes of "good" passengers, such as business travelers and tourists
- Passengers who do not fit into any of the clusters of "good" passengers will be subject to additional screening
- We are using Autoclass, a clustering system developed at NASA Ames




2/3/2004





## Preliminary Data Mining Results on Northwest Airlines data

- Currently we have one day's worth of NWA passenger data
  - It currently includes name and flight information
- We ran Autoclass, a Bayesian clustering system developed at NASA Ames, on this NWA data
  - It rediscovered the NWA hubs in Detroit, Minneapolis, Amsterdam, and Memphis, and discovered a cluster of passengers who appear to have been purged from the system



2/3/2004





## Preliminary Data Mining Results on LexisNexis data

- LexisNexis has given us access to their Web-based database of 180 million Americans, which includes previous addresses and phone numbers, license information, real estate information, etc.
- We assembled a database of pseudo-passengers, and used Voquette software to retrieve additional data on them from the LexisNexis Web site
- We ran GritBot on this database of pseudo-passengers and their LexisNexis data. Some of the anomalies it found include:
  - Two people who each paid \$100 for their homes
  - Someone who owns 12 homes



2/3/2004



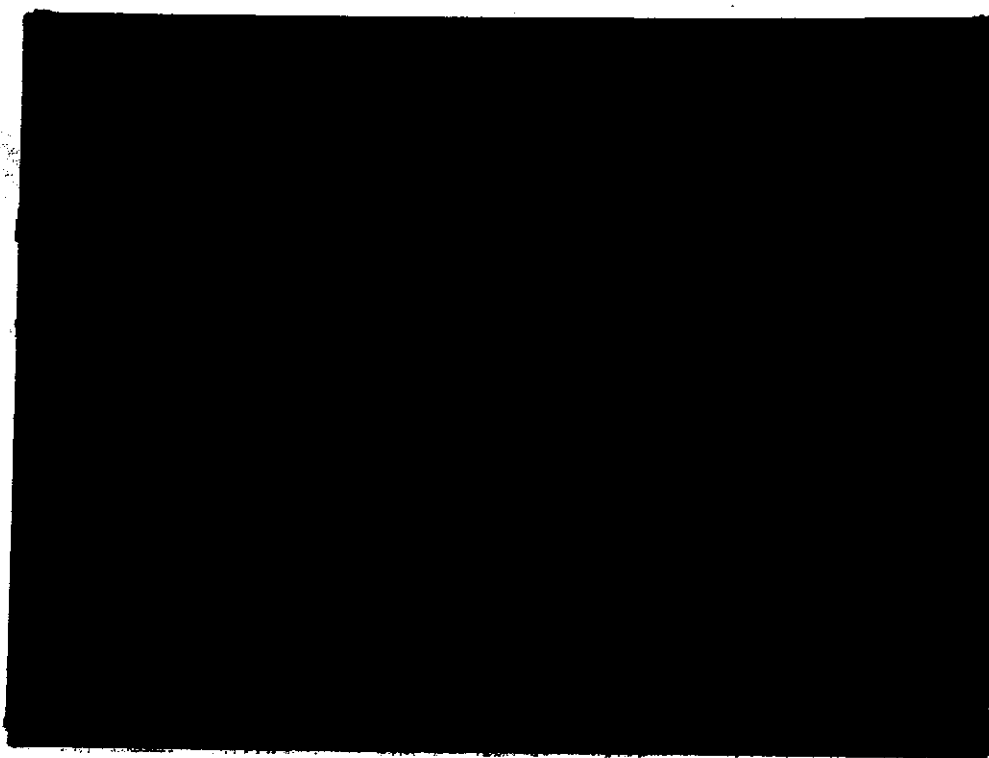
## Concern for Scalability

- Recognized that any effective system must be scalable to support
  - All 430 airports
  - 600 million passengers per year, with a potential growth to one billion passengers
- Looking at highly scalable computational infrastructures such as
  - Grid technology, with its evolving international standards and increasing commercial involvement
    - NASA's Information Power Grid, which is being targeted to support various NASA programs



2/3/2004





## Applicable NASA Core Technology

- **Data Mining and Analysis**
  - One of the top research labs in the country in statistical machine learning and unsupervised machine learning
  - Very close ties to broad academic and industrial research community
- **Distributed Heterogeneous Infrastructure**
  - Technology to access and organize data from multiple outside sources
  - Developing and fielding Information Power Grid (IPG) which is a seamless, scalable infrastructure integrating:
    - Distributed collections of processors, sensors and data
    - Uses what has become the de facto standard grid software





## Applicable NASA Core Technologies

- **Aviation Safety Reporting System (ASRS)**
  - Receives, processes, and analyses first-hand reports that describe unsafe occurrences and hazardous situations in the aviation system.
  - Uses formal data analysis methods to identify system risks and vulnerabilities.
- **Human Factors**
  - Twenty-five year record of providing the aviation industry with scientific principles and practical solutions for improving safety and effectiveness of crew training and line operations



2/3/2004





## Closing thoughts from the FAA's Third International Aviation Security Technology Symposium

- "Activity is no substitute for progress." [Richard Doney, from the UK Department of Transportation]
- "Don't depend too much on technology." [Joel Feldshul, former Israeli General (Chief of Intelligence), former chairman and CEO of El Al Airlines and now chairman and CEO of an Israeli Security Consulting firm.]
- "We need to ensure that we are not generals planning for the last war." [Nick Cartwright from Transport Canada]



2/3/2004

