Combining Voice and Vision for a Better User Experience

Ways Voice and Vision can Combine to Make Better Solutions

- **Biometric authentication**
  - Fusion of voice and vision
- **Demographic identification**
  - Looking at and hearing the person
    - Age, gender, ethnicity
- **Speech recognition**
  - M vs N: sounds similar, but looks very different
- **Emotions, scene analysis and other ID tasks**

Authentication: The Typical User Requirements

- **Accuracy – FA/FR and spoofing**
- **Must be Fast and Convenient**
- **Must Work in Real Environments**
  - Not carefully setup/crafted lab experiments
- **Must be Affordable**
  - Running on low cost platforms

Voice Triggers

- **Not Difficult to Develop a Voice Trigger Technology**
- **REALLY Difficult to Develop one that Works Well**
  - Noise, "normal" people, changing environments, distance, etc.
  - Low power – can't run down battery
  - User defined capability now a requirement
  - Low MIPS
  - Low Memory
- **Key – Balancing False Accept and False Reject**
  - You CANNOT test one without the other

Sensory Voice Triggers used in more than 1/2 Billion Products

Voice Password

- **Start with Sensory Voice Trigger Technology**
  - Fixed with enrollment
  - User Defined
- **Text Dependent**
  - User defines their own password through enrollment
    - Must be quick
  - System builds voice print
    - Must be local for speed, security and privacy
- **Text Independent**
  - User enrolls by speaking a paragraph
  - System build voice print based on users voice
    - Must be local for speed, security and privacy

sensory

Security

Authentication

- **Options Today**
  - PINs
  - Patterns
  - Finger Print
  - Iris Recognition
  - Voice Recognition
  - Face Recognition

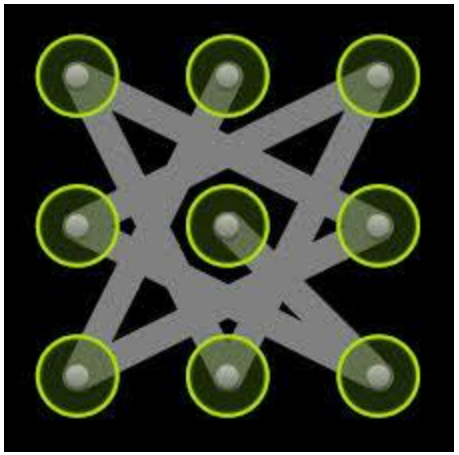
- **All of these can have issues when used in isolation**

Requirements for Authentication Solutions

- **Ease of Use**
  - Large percentage users don't lock their mobile devices. Why? Hassle.
  - If authentication is slow or challenging to use, people won't use
- **High Accuracy**
  - Must work and provide real security
- **Fast & Easy to Enroll/Setup**
- **Changeability**
- **Affordable**

Authentication PINS and Patterns

- **Issues**
  - If someone sees or learns your PIN/Pattern, then 100% breakable
  - Difficult to remember
  - Requires hands on device
  - Difficult to use while driving
  - Difficult to use while wearing gloves



Authentication
Iris

- **Issues**
  - Intrusive
  - Long enrollment
  - Doesn't work in poor lighting
  - If copied, hard to change
    - you only have two eyes
  - Nearly impossible to use (safely) while driving
  - Difficult to use while moving (walking, riding, running, etc.)

Authentication Finger Print

- **Issues with Finger Print**
  - Expensive, requires additional hardware
  - High rate of false rejects – frustrating to user
    - Dirty, greasy fingers, cuts, etc. all effect accuracy
  - If copied, hard to change
  - Easy to break
    - Tape, gummybears, glue, photo, etc.

Voice and Face

- **Issues with either as solo solution**
  - Voice
    - doesn't work in high noise
    - Voice changes over time or with colds, etc
  - Face
    - doesn't work in dark
    - Pitch, yah, angle all can effect
  - If done in the cloud, increased risk of lost personal data
- **Together Voice and Face can solve these issues**

Voice and Face for Authentication

- **Fusion of Voice and/or Vision**
  - Providing two biometrics
    - More environmentally robust
    - More accurate across all conditions
    - Ability to improve convenience
  - Allow the user to choose what level of security they want
    - For highest security, use FACE + VOICE = TrulySecure
  - Allow the user to choose what interface they want
    - For some features, face is fast and easy and provides the security desired

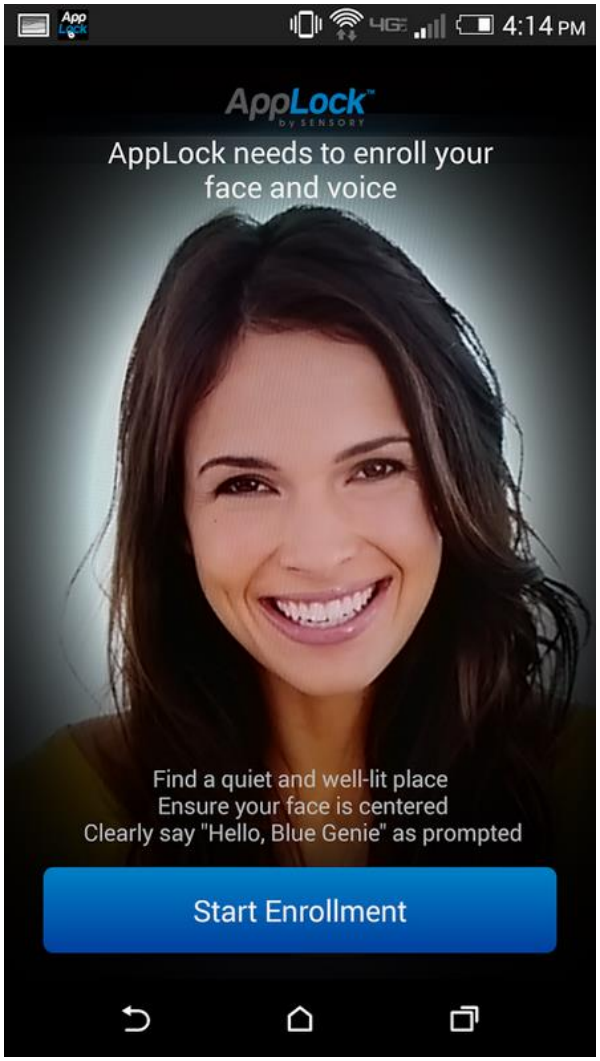
TrulySecure

- **Easy to Use and Enroll**
  - Fast enrollment <20 seconds
  - Fast authentication <2 seconds
  - NO need to roll eyes, move your face around from side to side, up and down
  - On mobile device, one hand ultra convenient enroll and use

TrulySecure

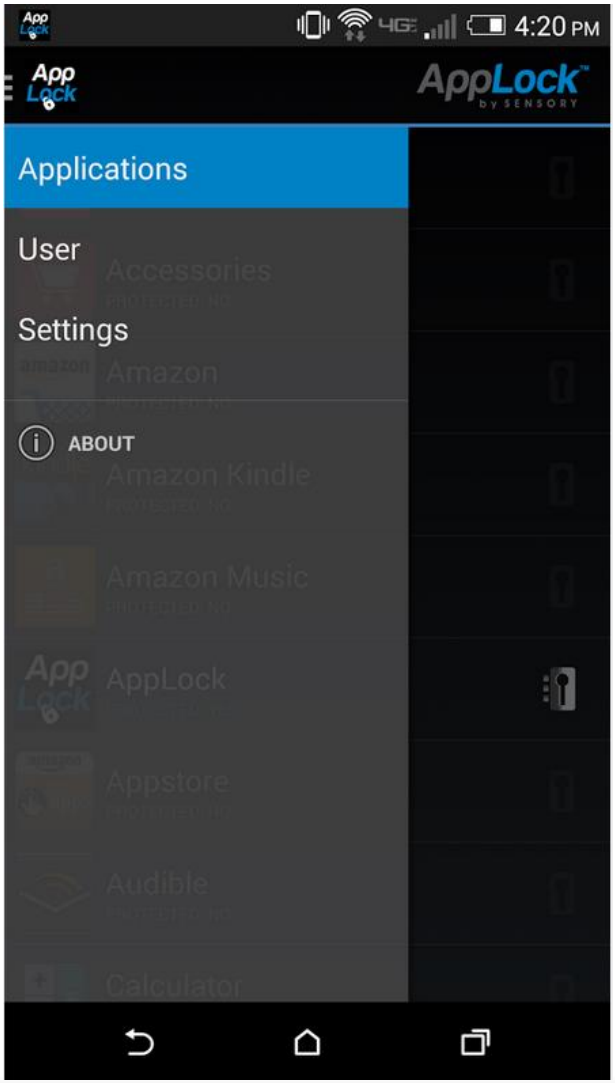
- **Adaptability**
  - Face recognition needs to be able to adapt and therefore improve
    - Glasses, sunglasses, strange angles, etc.
    - With each adapt, overall use improves and FR's reduce
- **Changeability**
  - Biometric must be replaceable if lost or stolen
- **Affordable**
  - On mobile, no additional/new hardware required
- **Secure**
  - FIDO compliant
  - See for yourself – search Applock by Sensory on Playstore

AppLock
Example of TrulySecure

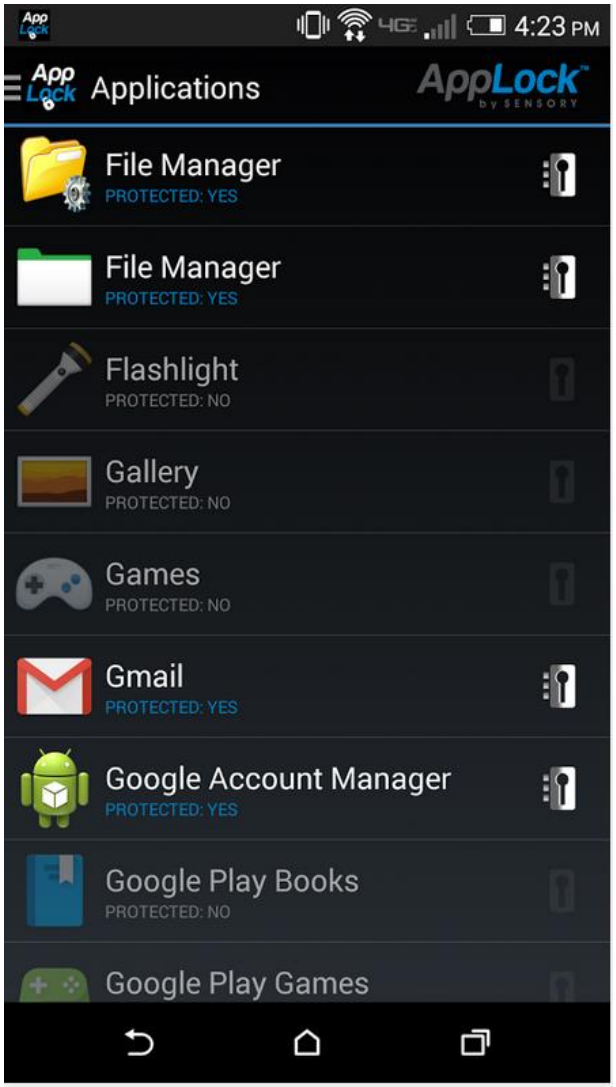


- **Enroll face and voice at the same time**
- **User can enroll with a fixed passphrase or user defined passphrase**

Applock

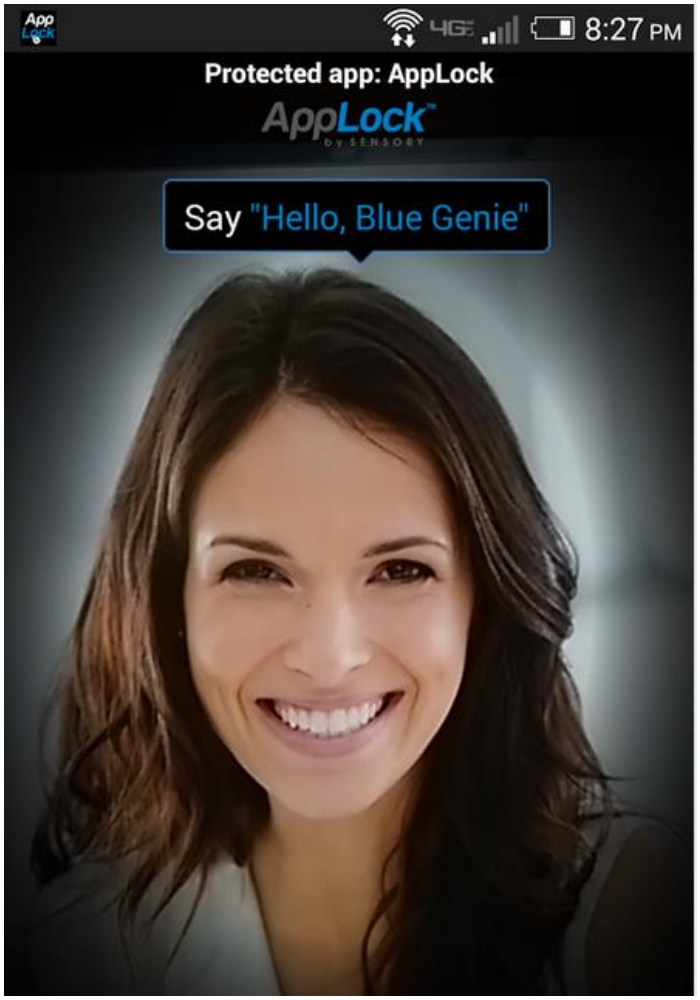


User chooses what apps they want to lock



Applock

- **User speaks voice password, or just looks at device, or looks at and speaks voice password for authentication**



TrulySecure

Complete Solution

- **TrulyHandsfree™ voice control**
  - Low power high accuracy voice triggers
  - Noise tolerant & flexible command sets
  - Speaker verification and identification
  - TrulyHandsfree 4.0 scheduled for Q3 2015

- **TrulySecure™ authentication**
  - Biometric fusion of face and voice
  - Accuracy with convenience
  - No specialized hardware required
  - TrulySecure 2.0 scheduled for Q2 2015

- **TrulyNatural™ fluent speech engine**
  - State of the art embedded deep net
  - Highest accuracy large vocabulary embedded
  - Flexibility in size and features
  - Capable of natural language interactions

Brand New!


Спасибо
Gracias
شكرا
Obrigado
Спасибо
Dank U
Grazie
Ευχαριστώ
Danke
Merci
Obrigado
Diolch
Thank You
Ngiyabonga
Dank U
Thank
You
Merci
Ngiyabonga
Tack
Ευχαριστώ
Dziękuję
Dziękuję
Danke
Grazie
תודה
Diolch
Terima Kasih
Diolch
Grazie
Tack
Ευχαριστώ
Tack
Diolch
Gracias
Merci
תודה