Active Linguistic Authentication Revisited

Real-Time Stylometric Evaluation towards Multi-Modal Decision Fusion

Ariel Stolerman † Alex Fridman † Rachel Greenstadt †
Patrick Brennan ‡ Patrick Juola ‡

†Drexel University
Philadelphia, PA

‡Juola & Associates
Pittsburgh, PA

10^{th} IFIP WG 11.9, January 2014

Outline

Outline

Introduction

- Active Authentication:
  - The process of continuously verifying a user based on his/her ongoing interaction with the computer
- Stylometry:
  - The study of linguistic style applied to user identification
- This work: Evaluating stylometric sensors for active authentication
  - Realtime-like multi-user environment
  - Time-based overlapping sliding windows ⇒
  - High-paced authentication decision making

Introduction

- Active Authentication:
  - The process of continuously verifying a user based on his/her ongoing interaction with the computer
- Stylometry:
  - The study of linguistic style applied to user identification
- This work: Evaluating stylometric sensors for active authentication
  - Realtime-like multi-user environment
  - Time-based overlapping sliding windows ⇒
  - High-paced authentication decision making

Introduction

- Active Authentication:
  - The process of continuously verifying a user based on his/her ongoing interaction with the computer
- Stylometry:
  - The study of linguistic style applied to user identification
- This work: Evaluating stylometric sensors for active authentication
  - Realtime-like multi-user environment
  - Time-based overlapping sliding windows ⇒
  - High-paced authentication decision making

Introduction

- Active Authentication:
  - The process of continuously verifying a user based on his/her ongoing interaction with the computer
- Stylometry:
  - The study of linguistic style applied to user identification
- This work: Evaluating stylometric sensors for active authentication
  - Realtime-like multi-user environment
  - Time-based overlapping sliding windows ⇒
  - High-paced authentication decision making

Introduction

- Active Authentication:
  - The process of continuously verifying a user based on his/her ongoing interaction with the computer
- Stylometry:
  - The study of linguistic style applied to user identification
- This work: Evaluating stylometric sensors for active authentication
  - Realtime-like multi-user environment
  - Time-based overlapping sliding windows $\Rightarrow$
  - High-paced authentication decision making

Introduction

- Active Authentication:
  - The process of continuously verifying a user based on his/her ongoing interaction with the computer
- Stylometry:
  - The study of linguistic style applied to user identification
- This work: Evaluating stylometric sensors for active authentication
  - Realtime-like multi-user environment
  - Time-based overlapping sliding windows $\Rightarrow$
  - High-paced authentication decision making

Motivation

- Increasing interest in behavioral biometrics [AT07]
  - Use common hardware to authenticate (mouse, keyboard)
- Yet achieved performance is mixed ⇒ better solutions required
- Most work uses static data
  - Active Linguistic Authentication Dataset [JNS+13]: typical HCI
- Stylometry: effective identifier
  - High-level behavioral biometric for authentication systems
- Primary goal: use stylometric sensors in multi-modal systems

`[ keyboard | mouse | web-browsing | stylometry ]`

Motivation

- Increasing interest in behavioral biometrics [AT07]
  - Use common hardware to authenticate (mouse, keyboard)
- Yet achieved performance is mixed ⇒ better solutions required
- Most work uses static data
  - Active Linguistic Authentication Dataset [JNS+13]: typical HCI
- Stylometry: effective identifier
  - High-level behavioral biometric for authentication systems
- Primary goal: use stylometric sensors in multi-modal systems

`[ keyboard | mouse | web-browsing | stylometry ]`

Motivation

- Increasing interest in behavioral biometrics [AT07]
  - Use common hardware to authenticate (mouse, keyboard)
- Yet achieved performance is mixed $\Rightarrow$ better solutions required
- Most work uses static data
  - Active Linguistic Authentication Dataset [JNS+13]: typical HCI
- Stylometry: effective identifier
  - High-level behavioral biometric for authentication systems
- Primary goal: use stylometric sensors in multi-modal systems

`[ keyboard | mouse | web-browsing | stylometry ]`

Motivation

- Increasing interest in behavioral biometrics [AT07]
  - Use common hardware to authenticate (mouse, keyboard)
- Yet achieved performance is mixed ⇒ better solutions required
- Most work uses static data
  - Active Linguistic Authentication Dataset [JNS+13]: typical HCI
- Stylometry: effective identifier
  - High-level behavioral biometric for authentication systems
- Primary goal: use stylometric sensors in multi-modal systems

`[ keyboard | mouse | web-browsing | stylometry ]`

Motivation

- Increasing interest in behavioral biometrics [AT07]
  - Use common hardware to authenticate (mouse, keyboard)
- Yet achieved performance is mixed ⇒ better solutions required
- Most work uses static data
  - Active Linguistic Authentication Dataset [JNS+13]: typical HCI
- Stylometry: effective identifier
  - High-level behavioral biometric for authentication systems
- Primary goal: use stylometric sensors in multi-modal systems

`[ keyboard | mouse | web-browsing | stylometry ]`

Motivation

- Increasing interest in behavioral biometrics [AT07]
  - Use common hardware to authenticate (mouse, keyboard)
- Yet achieved performance is mixed $\Rightarrow$ better solutions required
- Most work uses static data
  - Active Linguistic Authentication Dataset [JNS+13]: typical HCI
- Stylometry: effective identifier
  - High-level behavioral biometric for authentication systems
- Primary goal: use stylometric sensors in multi-modal systems

`[ keyboard | mouse | web-browsing | stylometry ]`

Motivation

- Increasing interest in behavioral biometrics [AT07]
  - Use common hardware to authenticate (mouse, keyboard)
- Yet achieved performance is mixed ⇒ better solutions required
- Most work uses static data
  - Active Linguistic Authentication Dataset [JNS+13]: typical HCI
- Stylometry: effective identifier
  - High-level behavioral biometric for authentication systems
- Primary goal: use stylometric sensors in multi-modal systems

`[ keyboard | mouse | web-browsing | stylometry ]`

Motivation – Contd.

- Implications on forensic: post-mortem stylometric analysis
  - Effective features in "noisy" environment?
  - Sliding window size, overlap?
  - Idle periods (no input)?

Motivation – Contd.

- Implications on forensic: post-mortem stylometric analysis
  - Effective features in “noisy” environment?
  - Sliding window size, overlap?
  - Idle periods (no input)?

Motivation – Contd.

- Implications on forensic: post-mortem stylometric analysis
  - Effective features in “noisy” environment?
  - Sliding window size, overlap?
  - Idle periods (no input)?

Motivation – Contd.

- Implications on forensic: post-mortem stylometric analysis
  - Effective features in “noisy” environment?
  - Sliding window size, overlap?
  - Idle periods (no input)?

Outline

Active Authentication

- Active authentication ⇒ sensor data varies with time
  - Verification must be on *recent* data only
- Different biometrics for different user activities
  - Web browsing ⇒ mouse + web activity
  - Document editing ⇒ keystrokes + stylometry
- ⇒ Multi-modal authentication [SZJK07]
  - Robust to dynamic real-time HCI
  - Sensor fusion ⇒ greater error reduction [AP95]

Active Authentication

- Active authentication $\Rightarrow$ sensor data varies with time
  - Verification must be on *recent* data only
- Different biometrics for different user activities
  - Web browsing $\Rightarrow$ mouse + web activity
  - Document editing $\Rightarrow$ keystrokes + stylometry
- $\Rightarrow$ Multi-modal authentication [SZJK07]
  - Robust to dynamic real-time HCI
  - Sensor fusion $\Rightarrow$ greater error reduction [AP95]

Active Authentication

- Active authentication ⇒ sensor data varies with time
  - Verification must be on *recent* data only
- Different biometrics for different user activities
  - Web browsing ⇒ mouse + web activity
  - Document editing ⇒ keystrokes + stylometry
- ⇒ Multi-modal authentication [SZJK07]
  - Robust to dynamic real-time HCI
  - Sensor fusion ⇒ greater error reduction [AP95]

Active Authentication

- Active authentication ⇒ sensor data varies with time
  - Verification must be on *recent* data only
- Different biometrics for different user activities
  - Web browsing ⇒ mouse + web activity
  - Document editing ⇒ keystrokes + stylometry
- ⇒ Multi-modal authentication [SZJK07]
  - Robust to dynamic real-time HCI
  - Sensor fusion ⇒ greater error reduction [AP95]

Active Authentication

- Active authentication ⇒ sensor data varies with time
  - Verification must be on *recent* data only
- Different biometrics for different user activities
  - Web browsing ⇒ mouse + web activity
  - Document editing ⇒ keystrokes + stylometry
- ⇒ Multi-modal authentication [SZJK07]
  - Robust to dynamic real-time HCI
  - Sensor fusion ⇒ greater error reduction [AP95]

Active Authentication

- Active authentication $\Rightarrow$ sensor data varies with time
  - Verification must be on *recent* data only
- Different biometrics for different user activities
  - Web browsing $\Rightarrow$ mouse + web activity
  - Document editing $\Rightarrow$ keystrokes + stylometry
- $\Rightarrow$ Multi-modal authentication [SZJK07]
  - Robust to dynamic real-time HCI
  - Sensor fusion $\Rightarrow$ greater error reduction [AP95]

Active Authentication

- Active authentication $\Rightarrow$ sensor data varies with time
  - Verification must be on *recent* data only
- Different biometrics for different user activities
  - Web browsing $\Rightarrow$ mouse + web activity
  - Document editing $\Rightarrow$ keystrokes + stylometry
- $\Rightarrow$ Multi-modal authentication [SZJK07]
  - Robust to dynamic real-time HCI
  - Sensor fusion $\Rightarrow$ greater error reduction [AP95]

Stylometry

- Stylometry: authorship attribution based on linguistic style
  - Features: function words, grammar, *n*-gram frequencies...
  - Used also for profiling: gender, age, native language
- In active authentication context: verification
  - Unary author-specific classifiers to determine `user/attacker`
- Challenges:
  - Open-world settings: much harder than “standard” stylometry
  - Inconsistent input frequency ⇔ stylometry data requirements
- But:
  - Unique idiosyncrasies like misspellings / keystroke patterns

Stylometry

- Stylometry: authorship attribution based on linguistic style
  - Features: function words, grammar, *n*-gram frequencies...
  - Used also for profiling: gender, age, native language
- In active authentication context: verification
  - Unary author-specific classifiers to determine `user/attacker`
- Challenges:
  - Open-world settings: much harder than "standard" stylometry
  - Inconsistent input frequency ⇔ stylometry data requirements
- But:
  - Unique idiosyncrasies like misspellings / keystroke patterns

Stylometry

- Stylometry: authorship attribution based on linguistic style
  - Features: function words, grammar, *n*-gram frequencies...
  - Used also for profiling: gender, age, native language
- In active authentication context: verification
  - Unary author-specific classifiers to determine `user/attacker`
- Challenges:
  - Open-world settings: much harder than "standard" stylometry
  - Inconsistent input frequency ⇔ stylometry data requirements
- But:
  - Unique idiosyncrasies like misspellings / keystroke patterns

Stylometry

- Stylometry: authorship attribution based on linguistic style
  - Features: function words, grammar, *n*-gram frequencies...
  - Used also for profiling: gender, age, native language
- In active authentication context: verification
  - Unary author-specific classifiers to determine `user/attacker`
- Challenges:
  - Open-world settings: much harder than “standard” stylometry
  - Inconsistent input frequency $\Leftrightarrow$ stylometry data requirements
- But:
  - Unique idiosyncrasies like misspellings / keystroke patterns

Stylometry

- Stylometry: authorship attribution based on linguistic style
  - Features: function words, grammar, *n*-gram frequencies...
  - Used also for profiling: gender, age, native language
- In active authentication context: verification
  - Unary author-specific classifiers to determine `user/attacker`
- Challenges:
  - Open-world settings: much harder than “standard” stylometry
  - Inconsistent input frequency $\Leftrightarrow$ stylometry data requirements
- But:
  - Unique idiosyncrasies like misspellings / keystroke patterns

Stylometry

- Stylometry: authorship attribution based on linguistic style
  - Features: function words, grammar, *n*-gram frequencies...
  - Used also for profiling: gender, age, native language
- In active authentication context: verification
  - Unary author-specific classifiers to determine `user/attacker`
- Challenges:
  - Open-world settings: much harder than “standard” stylometry
  - Inconsistent input frequency $\Leftrightarrow$ stylometry data requirements
- But:
  - Unique idiosyncrasies like misspellings / keystroke patterns

Stylometry

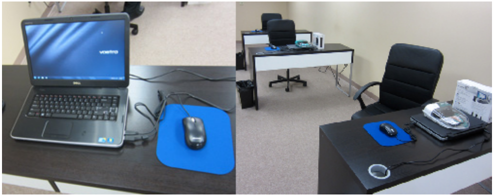
- Stylometry: authorship attribution based on linguistic style
  - Features: function words, grammar, *n*-gram frequencies...
  - Used also for profiling: gender, age, native language
- In active authentication context: verification
  - Unary author-specific classifiers to determine `user/attacker`
- Challenges:
  - Open-world settings: much harder than “standard” stylometry
  - Inconsistent input frequency ⇔ stylometry data requirements
- But:
  - Unique idiosyncrasies like misspellings / keystroke patterns

Outline

1. Motivation
2. Background
3. Dataset
4. Methodology
5. Evaluation
6. Conclusion

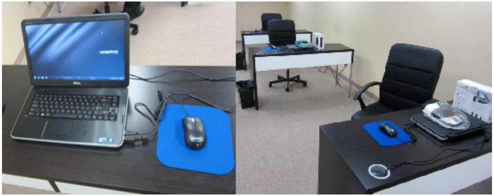
Dataset Collection

- Used the full Active Linguistic Authentication Dataset [JNS+13]
- Computer input collected in a simulated work environment
  - One 40-hour week data collected from 80 temps
  - Research and writings tasks
  - Tracked keyboard, mouse and web browsing behavior
  - Uniformity: same hardware for all workers



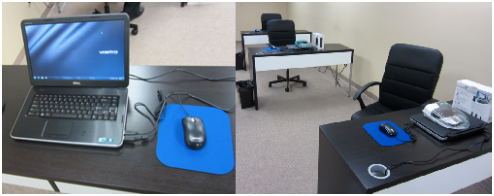
Dataset Collection

- Used the full Active Linguistic Authentication Dataset [JNS+13]
- Computer input collected in a simulated work environment
  - One 40-hour week data collected from 80 temps
  - Research and writings tasks
  - Tracked keyboard, mouse and web browsing behavior
  - Uniformity: same hardware for all workers



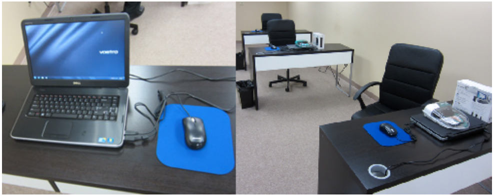
Dataset Collection

- Used the full Active Linguistic Authentication Dataset [JNS+13]
- Computer input collected in a simulated work environment
  - One 40-hour week data collected from 80 temps
  - Research and writings tasks
  - Tracked keyboard, mouse and web browsing behavior
  - Uniformity: same hardware for all workers



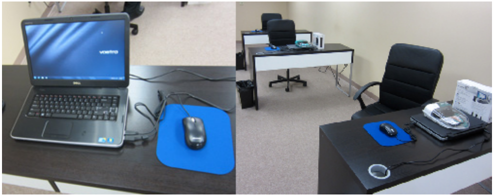
Dataset Collection

- Used the full Active Linguistic Authentication Dataset [JNS+13]
- Computer input collected in a simulated work environment
  - One 40-hour week data collected from 80 temps
  - Research and writings tasks
  - Tracked keyboard, mouse and web browsing behavior
  - Uniformity: same hardware for all workers



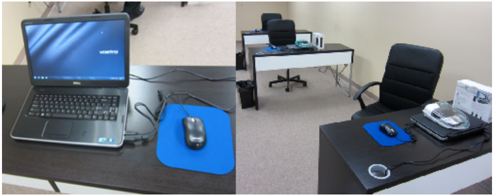
Dataset Collection

- Used the full Active Linguistic Authentication Dataset [JNS+13]
- Computer input collected in a simulated work environment
  - One 40-hour week data collected from 80 temps
  - Research and writings tasks
  - Tracked keyboard, mouse and web browsing behavior
  - Uniformity: same hardware for all workers



Dataset Collection

- Used the full Active Linguistic Authentication Dataset [JNS+13]
- Computer input collected in a simulated work environment
  - One 40-hour week data collected from 80 temps
  - Research and writings tasks
  - Tracked keyboard, mouse and web browsing behavior
  - Uniformity: same hardware for all workers



Dataset Preprocessing

- Used 67 of the 80 users w/ minimum 16.67 hours of activity

Min. chars/user	17,027
Max chars/user	263,165
AVg.	84,206
Total	5,641,788

- All 5-day keyboard and mouse events gathered in one file/user
  - Divided into 5 equal-size folds
  - Reduced all > 2 min idle periods to *exactly* 2 min
  - Collected *all* keystrokes: alphanumeric & special keys (`shift, ctrl, backspace...`)
  - Special chars converted to 1-char placeholders (e.g. `backspace` $\Rightarrow$ β)

Dataset Preprocessing

- Used 67 of the 80 users w/ minimum 16.67 hours of activity

Min. chars/user	17,027
Max chars/user	263,165
AVg.	84,206
Total	5,641,788

- All 5-day keyboard and mouse events gathered in one file/user
  - Divided into 5 equal-size folds
  - Reduced all > 2 min idle periods to *exactly* 2 min
  - Collected *all* keystrokes: alphanumeric & special keys (`shift, ctrl, backspace...`)
  - Special chars converted to 1-char placeholders (e.g. `backspace` $\Rightarrow$ β)

Dataset Preprocessing

- Used 67 of the 80 users w/ minimum 16.67 hours of activity

Min. chars/user	17,027
Max chars/user	263,165
AVg.	84,206
Total	5,641,788

- All 5-day keyboard and mouse events gathered in one file/user
  - Divided into 5 equal-size folds
  - Reduced all > 2 min idle periods to *exactly* 2 min
  - Collected *all* keystrokes: alphanumeric & special keys (`shift, ctrl, backspace...`)
  - Special chars converted to 1-char placeholders (e.g. `backspace` $\Rightarrow$ β)

Dataset Preprocessing

- Used 67 of the 80 users w/ minimum 16.67 hours of activity

Min. chars/user	17,027
Max chars/user	263,165
AVg.	84,206
Total	5,641,788

- All 5-day keyboard and mouse events gathered in one file/user
  - Divided into 5 equal-size folds
  - Reduced all > 2 min idle periods to *exactly* 2 min
  - Collected *all* keystrokes: alphanumeric & special keys (`shift, ctrl, backspace...`)
  - Special chars converted to 1-char placeholders (e.g. `backspace` $\Rightarrow$ β)

Dataset Preprocessing

- Used 67 of the 80 users w/ minimum 16.67 hours of activity

Min. chars/user	17,027
Max chars/user	263,165
AVg.	84,206
Total	5,641,788

- All 5-day keyboard and mouse events gathered in one file/user
  - Divided into 5 equal-size folds
  - Reduced all > 2 min idle periods to *exactly* 2 min
  - Collected *all* keystrokes: alphanumeric & special keys (`shift, ctrl, backspace...`)
  - Special chars converted to 1-char placeholders (e.g. `backspace` $\Rightarrow$ β)

Dataset Preprocessing

- Used 67 of the 80 users w/ minimum 16.67 hours of activity

Min. chars/user	17,027
Max chars/user	263,165
AVg.	84,206
Total	5,641,788

- All 5-day keyboard and mouse events gathered in one file/user
  - Divided into 5 equal-size folds
  - Reduced all > 2 min idle periods to *exactly* 2 min
  - Collected *all* keystrokes: alphanumeric & special keys (`shift, ctrl, backspace...`)
  - Special chars converted to 1-char placeholders (e.g. `backspace` $\Rightarrow$ β)

Outline

Challenges and Limitations

- Potential performance issues
  - On-the-fly heavy linguistic processing
- Authenticating input not designated for authentication
  - Credentials collection
  - Secure processing & storage is required

Challenges and Limitations

- Potential performance issues
  - On-the-fly heavy linguistic processing
- Authenticating input not designated for authentication
  - Credentials collection
  - Secure processing & storage is required

Challenges and Limitations

- Potential performance issues
  - On-the-fly heavy linguistic processing
- Authenticating input not designated for authentication
  - Credentials collection
  - Secure processing & storage is required

Challenges and Limitations

- Potential performance issues
  - On-the-fly heavy linguistic processing
- Authenticating input not designated for authentication
  - Credentials collection
  - Secure processing & storage is required

Challenges and Limitations

- Potential performance issues
  - On-the-fly heavy linguistic processing
- Authenticating input not designated for authentication
  - Credentials collection
  - Secure processing & storage is required

Previous Evaluation

- Initial data-based windows analysis on 14 users [JNS$^+$13]:
  - Day windows: 88% accuracy
    k-NN classifier + Manhattan distance + char n-grams
  - 100-1000-word windows: 93% accuracy
    No overlap, SVM classifier + extensive feature set
- Proof of concept, but:
  - Only 14 users
  - Data-wise windows (whole day, X words) – not useful for AA systems
    - Fast decision making is required
    - Miss attacks ("bad" windows w/ "good" data)

Previous Evaluation

- Initial data-based windows analysis on 14 users [JNS+13]:
  - Day windows: 88% accuracy
    k-NN classifier + Manhattan distance + char n-grams
  - 100-1000-word windows: 93% accuracy
    No overlap, SVM classifier + extensive feature set
- Proof of concept, but:
  - Only 14 users
  - Data-wise windows (whole day, X words) – not useful for AA systems
    - Fast decision making is required
    - Miss attacks ("bad" windows w/ "good" data)

Previous Evaluation

- Initial data-based windows analysis on 14 users [JNS+13]:
  - Day windows: 88% accuracy
    k-NN classifier + Manhattan distance + char n-grams
  - 100-1000-word windows: 93% accuracy
    No overlap, SVM classifier + extensive feature set
- Proof of concept, but:
  - Only 14 users
  - Data-wise windows (whole day, X words) – not useful for AA systems
    - Fast decision making is required
    - Miss attacks ("bad" windows w/ "good" data)

Previous Evaluation

- Initial data-based windows analysis on 14 users [JNS+13]:
  - Day windows: 88% accuracy
    k-NN classifier + Manhattan distance + char n-grams
  - 100-1000-word windows: 93% accuracy
    No overlap, SVM classifier + extensive feature set
- Proof of concept, but:
  - Only 14 users
  - Data-wise windows (whole day, X words) – not useful for AA systems
    - Fast decision making is required
    - Miss attacks ("bad" windows w/ "good" data)

Previous Evaluation

- Initial data-based windows analysis on 14 users [JNS$^+$13]:
  - Day windows: 88% accuracy
    k-NN classifier + Manhattan distance + char n-grams
  - 100-1000-word windows: 93% accuracy
    No overlap, SVM classifier + extensive feature set
- Proof of concept, but:
  - Only 14 users
  - Data-wise windows (whole day, X words) – not useful for AA systems
    - Fast decision making is required
    - Miss attacks (“bad” windows w/ “good” data)

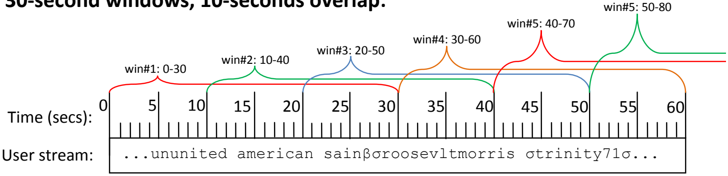
Previous Evaluation

- Initial data-based windows analysis on 14 users [JNS$^+$13]:
  - Day windows: 88% accuracy
    k-NN classifier + Manhattan distance + char *n*-grams
  - 100-1000-word windows: 93% accuracy
    No overlap, SVM classifier + extensive feature set
- Proof of concept, but:
  - Only 14 users
  - Data-wise windows (whole day, *X* words) – not useful for AA systems
    - Fast decision making is required
    - Miss attacks ("bad" windows w/ "good" data)

Real-Time Approach

- Closed-world classifier: one SVM trained on all 67 users
- Time-wise overlapping windows:
  - Size ($\geq$): 10, 30, 60, 300, 600 & 1200 secs
  - Overlap ($\leq$): 10 / 60 secs
- Overlap: allow sufficient data + frequent decisions
- Elimination of idle periods allows maximum #windows evaluation

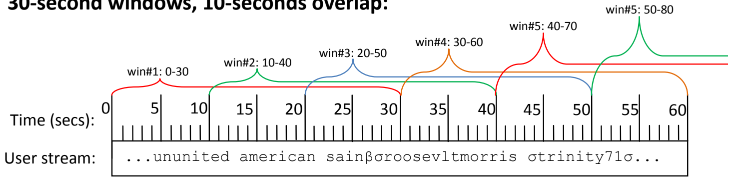
30-second windows, 10-seconds overlap:



Real-Time Approach

- Closed-world classifier: one SVM trained on all 67 users
- Time-wise overlapping windows:
  - Size ($\geq$): 10, 30, 60, 300, 600 & 1200 secs
  - Overlap ($\leq$): 10 / 60 secs
- Overlap: allow sufficient data + frequent decisions
- Elimination of idle periods allows maximum #windows evaluation

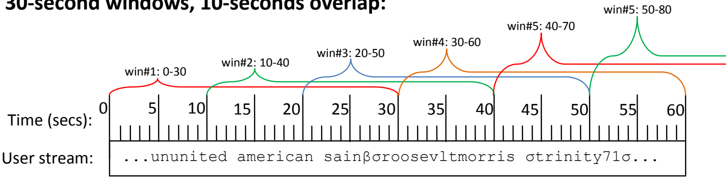
30-second windows, 10-seconds overlap:



Real-Time Approach

- Closed-world classifier: one SVM trained on all 67 users
- Time-wise overlapping windows:
  - Size ($\geq$): 10, 30, 60, 300, 600 & 1200 secs
  - Overlap ($\leq$): 10 / 60 secs
- Overlap: allow sufficient data + frequent decisions
- Elimination of idle periods allows maximum #windows evaluation

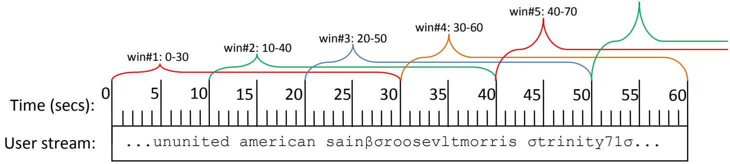
30-second windows, 10-seconds overlap:



Real-Time Approach

- Closed-world classifier: one SVM trained on all 67 users
- Time-wise overlapping windows:
  - Size ($\geq$): 10, 30, 60, 300, 600 & 1200 secs
  - Overlap ($\leq$): 10 / 60 secs
- Overlap: allow sufficient data + frequent decisions
- Elimination of idle periods allows maximum #windows evaluation

30-second windows, 10-seconds overlap:



Real-Time Approach – Contd.

- The AA feature-set:
  - Variation of the Writeprints feature-set [AC08]
  - Vast range of linguistic features

    char/word/POS *n*-grams, function words, word lengths, digits
  - "Applies" special chars: $\texttt{ch}\beta\beta\texttt{Cch}\beta\beta\texttt{hicago} \Rightarrow \texttt{Chicago}$
  - Frequency-based features normalization
  - Extracted using the JStylo authorship attribution framework [MAC+12]

Real-Time Approach – Contd.

- The AA feature-set:
  - Variation of the Writeprints feature-set [AC08]
  - Vast range of linguistic features
    char/word/POS *n*-grams, function words, word lengths, digits
  - "Applies" special chars: `chββCchββhicago` ⇒ `Chicago`
  - Frequency-based features normalization
  - Extracted using the JStylo authorship attribution framework [MAC+12]

Real-Time Approach – Contd.

- The AA feature-set:
  - Variation of the Writeprints feature-set [AC08]
  - Vast range of linguistic features
    char/word/POS *n*-grams, function words, word lengths, digits
  - “Applies” special chars: $\mathtt{ch}\beta\beta\mathtt{Cch}\beta\beta\mathtt{hicago} \Rightarrow \mathtt{Chicago}$
  - Frequency-based features normalization
  - Extracted using the JStylo authorship attribution framework [MAC+12]

Real-Time Approach – Contd.

- The AA feature-set:
  - Variation of the Writeprints feature-set [AC08]
  - Vast range of linguistic features
    char/word/POS *n*-grams, function words, word lengths, digits
  - "Applies" special chars: `chββCchββhicago` ⇒ `Chicago`
  - Frequency-based features normalization
  - Extracted using the JStylo authorship attribution framework [MAC+12]

Real-Time Approach – Contd.

- The AA feature-set:
  - Variation of the Writeprints feature-set [AC08]
  - Vast range of linguistic features
    char/word/POS *n*-grams, function words, word lengths, digits
  - “Applies” special chars: $\texttt{ch}\beta\beta\texttt{Cch}\beta\beta\texttt{hicago} \Rightarrow \texttt{Chicago}$
  - Frequency-based features normalization
  - Extracted using the JStylo authorship attribution framework [MAC+12]

Real-Time Approach – Contd.

- The AA feature-set:
  - Variation of the Writeprints feature-set [AC08]
  - Vast range of linguistic features
    char/word/POS *n*-grams, function words, word lengths, digits
  - “Applies” special chars: $\texttt{ch}\beta\beta\texttt{Cch}\beta\beta\texttt{hicago} \Rightarrow \texttt{Chicago}$
  - Frequency-based features normalization
  - Extracted using the JStylo authorship attribution framework [MAC+12]

Real-Time Approach – Contd.

- Applied minimum chars/window thresholds
  - Discarded too small windows (below threshold): "not enough data to decide"
  - Thresholds: 100–1000 characters (steps of 100)
  - Availability ⇔ potential accuracy
- Only sensors w/ train data for *all* users after filtering are kept
  - 37 of the 60 sensor configurations are kept

Real-Time Approach – Contd.

- Applied minimum chars/window thresholds
  - Discarded too small windows (below threshold): "not enough data to decide"
  - Thresholds: 100–1000 characters (steps of 100)
  - Availability ⇔ potential accuracy
- Only sensors w/ train data for *all* users after filtering are kept
  - 37 of the 60 sensor configurations are kept

Real-Time Approach – Contd.

- Applied minimum chars/window thresholds
  - Discarded too small windows (below threshold): "not enough data to decide"
  - Thresholds: 100–1000 characters (steps of 100)
  - Availability ⇔ potential accuracy
- Only sensors w/ train data for *all* users after filtering are kept
  - 37 of the 60 sensor configurations are kept

Real-Time Approach – Contd.

- Applied minimum chars/window thresholds
  - Discarded too small windows (below threshold): “not enough data to decide”
  - Thresholds: 100–1000 characters (steps of 100)
  - Availability ⇔ potential accuracy
- Only sensors w/ train data for *all* users after filtering are kept
  - 37 of the 60 sensor configurations are kept

Real-Time Approach – Contd.

- Applied minimum chars/window thresholds
  - Discarded too small windows (below threshold): "not enough data to decide"
  - Thresholds: 100–1000 characters (steps of 100)
  - Availability ⇔ potential accuracy
- Only sensors w/ train data for *all* users after filtering are kept
  - 37 of the 60 sensor configurations are kept

Real-Time Approach – Contd.

- Applied minimum chars/window thresholds
  - Discarded too small windows (below threshold): "not enough data to decide"
  - Thresholds: 100–1000 characters (steps of 100)
  - Availability ⇔ potential accuracy
- Only sensors w/ train data for *all* users after filtering are kept
  - 37 of the 60 sensor configurations are kept

Outline

Evaluation

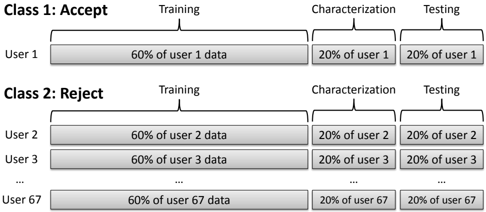
- Measure averaged FAR & FRR on 5-fold cyclic cross-validation
- Stylometry sensors intended for a multimodal system
  - Requires knowledge of expected FAR/FRR
- $\Rightarrow$ analysis technique:

Evaluation

- Measure averaged FAR & FRR on 5-fold cyclic cross-validation
- Stylometry sensors intended for a multimodal system
  - Requires knowledge of expected FAR/FRR
- $\Rightarrow$ analysis technique:

Evaluation

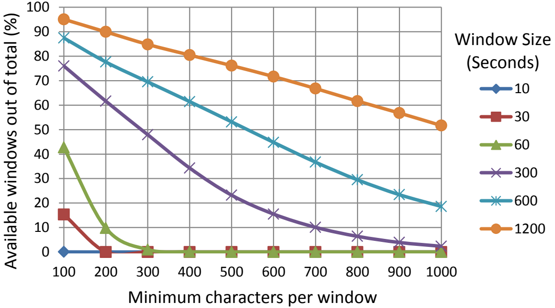
- Measure averaged FAR & FRR on 5-fold cyclic cross-validation
- Stylometry sensors intended for a multimodal system
    - Requires knowledge of expected FAR/FRR
- $\Rightarrow$ analysis technique:



Evaluation – Contd.

Availability by minimum char thresholds:

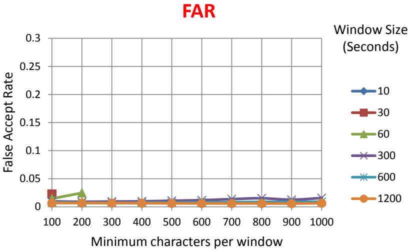
- The larger the window ⇒ the higher its availability
- Windows < 300 secs – not very useful

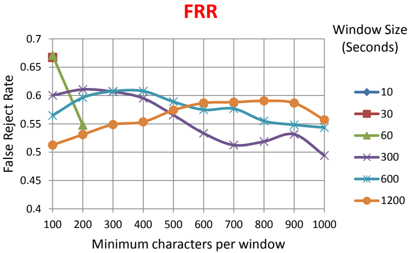


Evaluation – Contd.

Average FAR/FRR:

- Strict sensors
- The larger the window ⇒ the less affected by min char thresholds





Outline

Conclusion

- Proof of concept in [JNS+13] – insufficient for real-world settings
  - With small time-wise windows: performance deteriorates drastically
- Still, can be used in a mixture-of-experts approach – multi-modal systems
- Classification approach still limited
  - Should attempt open-world classifiers
  - Should attempt low-level linguistic features, typing patterns
- Immediate next step: fusion
  - Initial closed-world eval on 19 users: $<$ 1% FAR/FRR! [FSA+13]
  - To be continued...

Conclusion

- Proof of concept in [JNS+13] – insufficient for real-world settings
  - With small time-wise windows: performance deteriorates drastically
- Still, can be used in a mixture-of-experts approach – multi-modal systems
- Classification approach still limited
  - Should attempt open-world classifiers
  - Should attempt low-level linguistic features, typing patterns
- Immediate next step: fusion
  - Initial closed-world eval on 19 users: $<$ 1% FAR/FRR! [FSA+13]
  - To be continued...

Conclusion

- Proof of concept in [JNS+13] – insufficient for real-world settings
  - With small time-wise windows: performance deteriorates drastically
- Still, can be used in a mixture-of-experts approach – multi-modal systems
- Classification approach still limited
  - Should attempt open-world classifiers
  - Should attempt low-level linguistic features, typing patterns
- Immediate next step: fusion
  - Initial closed-world eval on 19 users: $<$ 1% FAR/FRR! [FSA+13]
  - To be continued...

Conclusion

- Proof of concept in [JNS$^+$13] – insufficient for real-world settings
  - With small time-wise windows: performance deteriorates drastically
- Still, can be used in a mixture-of-experts approach – multi-modal systems
- Classification approach still limited
  - Should attempt open-world classifiers
  - Should attempt low-level linguistic features, typing patterns
- Immediate next step: fusion
  - Initial closed-world eval on 19 users: $<$ 1% FAR/FRR! [FSA$^+$13]
  - To be continued...

Conclusion

- Proof of concept in [JNS+13] – insufficient for real-world settings
  - With small time-wise windows: performance deteriorates drastically
- Still, can be used in a mixture-of-experts approach – multi-modal systems
- Classification approach still limited
  - Should attempt open-world classifiers
  - Should attempt low-level linguistic features, typing patterns
- Immediate next step: fusion
  - Initial closed-world eval on 19 users: $<$ 1% FAR/FRR! [FSA+13]
  - To be continued...

Conclusion

- Proof of concept in [JNS+13] – insufficient for real-world settings
  - With small time-wise windows: performance deteriorates drastically
- Still, can be used in a mixture-of-experts approach – multi-modal systems
- Classification approach still limited
  - Should attempt open-world classifiers
  - Should attempt low-level linguistic features, typing patterns
- Immediate next step: fusion
  - Initial closed-world eval on 19 users: $<$ 1% FAR/FRR! [FSA+13]
  - To be continued...

Conclusion

- Proof of concept in [JNS+13] – insufficient for real-world settings
  - With small time-wise windows: performance deteriorates drastically
- Still, can be used in a mixture-of-experts approach – multi-modal systems
- Classification approach still limited
  - Should attempt open-world classifiers
  - Should attempt low-level linguistic features, typing patterns
- Immediate next step: fusion
  - Initial closed-world eval on 19 users: $<$ 1% FAR/FRR! [FSA+13]
  - To be continued...

Conclusion

- Proof of concept in [JNS+13] – insufficient for real-world settings
  - With small time-wise windows: performance deteriorates drastically
- Still, can be used in a mixture-of-experts approach – multi-modal systems
- Classification approach still limited
  - Should attempt open-world classifiers
  - Should attempt low-level linguistic features, typing patterns
- Immediate next step: fusion
  - Initial closed-world eval on 19 users: $<$ 1% FAR/FRR! [FSA+13]
  - To be continued...

Thank You

Thank You!

Questions?

- Contact: stolerman@cs.drexel.edu
- Drexel Privacy Security & Automation Lab: http://psal.cs.drexel.edu/
- Drexel Data Fusion Lab: http://dfl.ece.drexel.edu/
- Juola & Associates: http://juolaassociates.com/

For Further Reading I

Ahmed Abbasi and Hsinchun Chen.
Writeprints: A stylometric approach to identity-level identification and similarity detection in cyberspace.
ACM Trans. Inf. Syst., 26(2):1–29, 2008.

K.M. Ali and M.J. Pazzani.
On the link between error correlation and error reduction in decision tree ensembles.
Citeseer, 1995.

A.A.E. Ahmed and I. Traore.
A new biometric technology based on mouse dynamics.
Dependable and Secure Computing, IEEE Transactions on, 4(3):165 –179, july-sept. 2007.

Alex Fridman, Ariel Stolerman, Sayandeep Acharya, Patrick Brennan, Patrick Juola, Rachel Greenstadt, and Moshe Kam.
Decision fusion for multimodal active authentication.
IT Professional, 15(4):29–33, 2013.

Patrick Juola, John Noecker, Jr., Ariel Stolerman, Michael V. Ryan, Patrick Brennan, and Rachel Greenstadt.
A dataset for active linguistic authentication.
In *Proceedings of the Ninth Annual IFIP WG 11.9 International Conference on Digital Forensics*, Orlando, Florida, USA, January 2013. National Center for Forensic Science.

Andrew W. E. McDonald, Sadia Afroz, Aylin Caliskan, Ariel Stolerman, and Rachel Greenstadt.
Use fewer instances of the letter "i": Toward writing style anonymization.
In *Lecture Notes in Computer Science*, volume 7384, pages 299–318. Springer, 2012.

For Further Reading II

T. Sim, S. Zhang, R. Janakiraman, and S. Kumar.
Continuous verification using multimodal biometrics.
Pattern Analysis and Machine Intelligence, IEEE Transactions on, 29(4):687–700, 2007.