Cheating Detection: Identifying Fraud in Digital Exams

Bastian Küppers, Julia Opgen-Rhein, Thomas Eifert, Ulrik Schroeder
E-Assessment

Our Project: FLEX

FLEX (Framework for FLExible Electronic EXaminations)
Statement of the Problem (1 / 2)

- Cheating is a problem in examinations and can have many forms

- Electronic exams come with an apparently increased danger of impersonation and illegal communication between students

- This problem gets worse in a BYOD scenario
Stop. BYOD in an exam ???

Obstacles for E-Assessment (among others)

• In the exam weeks, almost all lecture rooms are in use
• Most of them have no permanent room setup with PCs/Laptops
• Setup per exam: large exam (~ 1000 students) requires handling of 1000 Laptops…

But:
• Apparently, every student has a laptop

⇒ BYOD appears as a way to enable E-Assessment on a broader base
Statement of the Problem (2 / 2)

• Existing solutions to security issues in Digital Examinations have multiple drawbacks for BYOD
  – Not guaranteed to be secure, as students’ devices are *untrusted platforms*
  – No available tool supports every major operating system

• A solution to secure Digital Exams in a BYOD setting has to be found
In-situ Attribution

- Monitor students’ during the exam for illicit activities, instead of locking the devices
  - Knowledge about possible cheating attempts has to be available to detect these activities
  - Particular cheating attempts may remain undetected

- To prevent plagiarism, the identity of the author of the examination’s results has to be determined
  - Student-related patterns in the log of events have to be identified
  - Typing patterns are a possible solution
A-posteriori Attribution (1 / 3)

- Analysis of the available log data produced during the Digital Exam
  - Interpretation as a time series

- Several techniques for analysis available
  - Process mining
  - Wavelet analysis
  - Author Verification
A-posteriori Attribution (2 / 3)

• Process mining
  – Used to discover processes, check conformance with a process model or improve existing processes
  – Assumption: cheating generates a different process model than regularly working on the exam’s assignments

• Wavelet / time series Analysis
  – Used to analyze linear time-frequency functions
  – The amount of answers that a student has entered into the system is interpreted as a frequency
    • High amount of answers relates to a high frequency
    • Low amount of answers relates to a low frequency
  – Assumption: The decomposition of the frequency signals reveals different frequencies for cheating
A-posteriori Attribution (3 / 3)

- For written texts and programming assignments, the submissions of the students can be compared with previous work from assignments and tutorials

- Previous material is used to learn the linguistic / programming style of a student

- This style is compared to the style that is inherent to the submission for the Digital Exam
Cheating Detection: Identifying Fraud in Digital Exams

Conditions

• A sufficient amount of data has to be available

• Therefore, not only final submission is monitored and analyzed, but also intermediate results during the exam, network activity…

• The data has to be available with a time stamp

• The collection of the data must not influence the performance of the students’ devices

• For author verification, reference material has to be collected during the semester
Conduction of Exams: Analogous vs. Digital

Summary

• Cheating detection for Digital Exams requires different measures than for paper-based exams

• Analyzing students’ submissions can only indicate a cheating attempt, but not prove it

➜ Combination of cheating prevention (during exam) and cheating detection (during / after exam) makes cheating attempts difficult and risky

• Current steps include the prototypical implementation of the proposed ways of a-posteriori cheating
Thanks for your attention! 😊
Takk for oppmerksomheten! 😊

Are there any questions or comments?