

Privacy Preserving Speech Processing



Isabel Trancoso
José Portêlo
INESC-ID/IST, Univ. Lisboa

Bhiksha Raj
LTI
Carnegie Mellon University

Gerard Chollet
Nigel Cannings
Intelligent Voice



Dijana Petrovska-Delacretaz
Télécom SudParis, CNRS

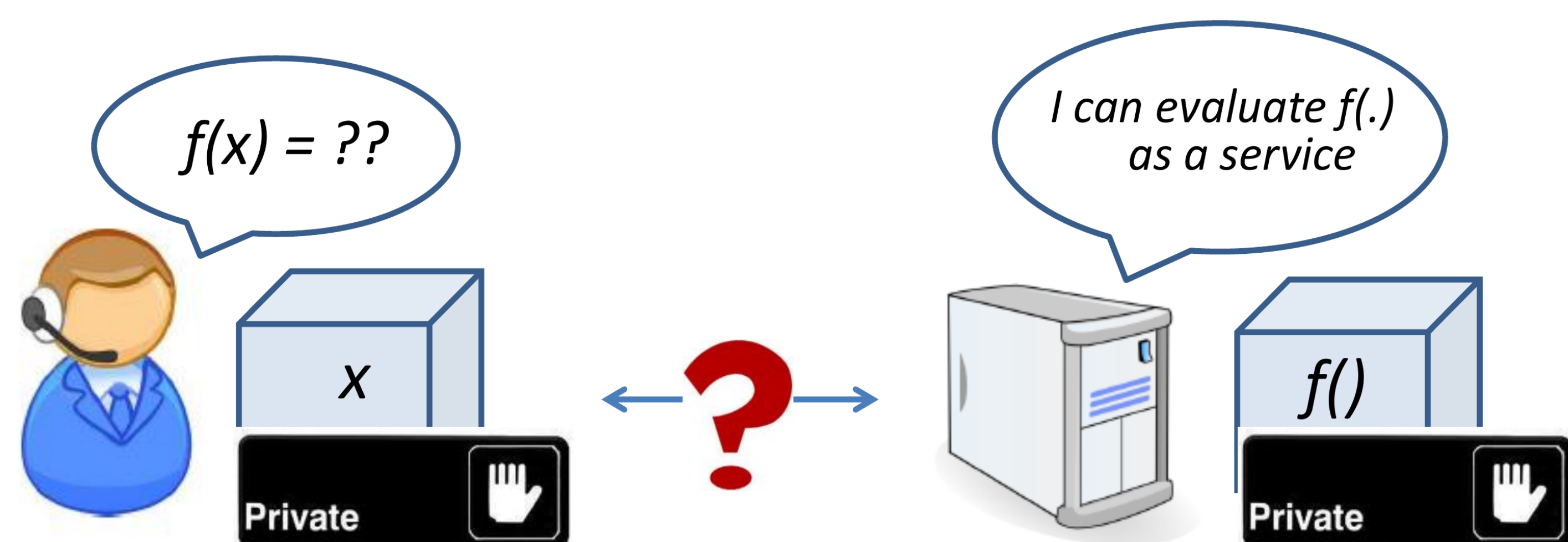
Atta Badii
University of Reading

Jean-Jacques Quisquater
UCL Crypto Group

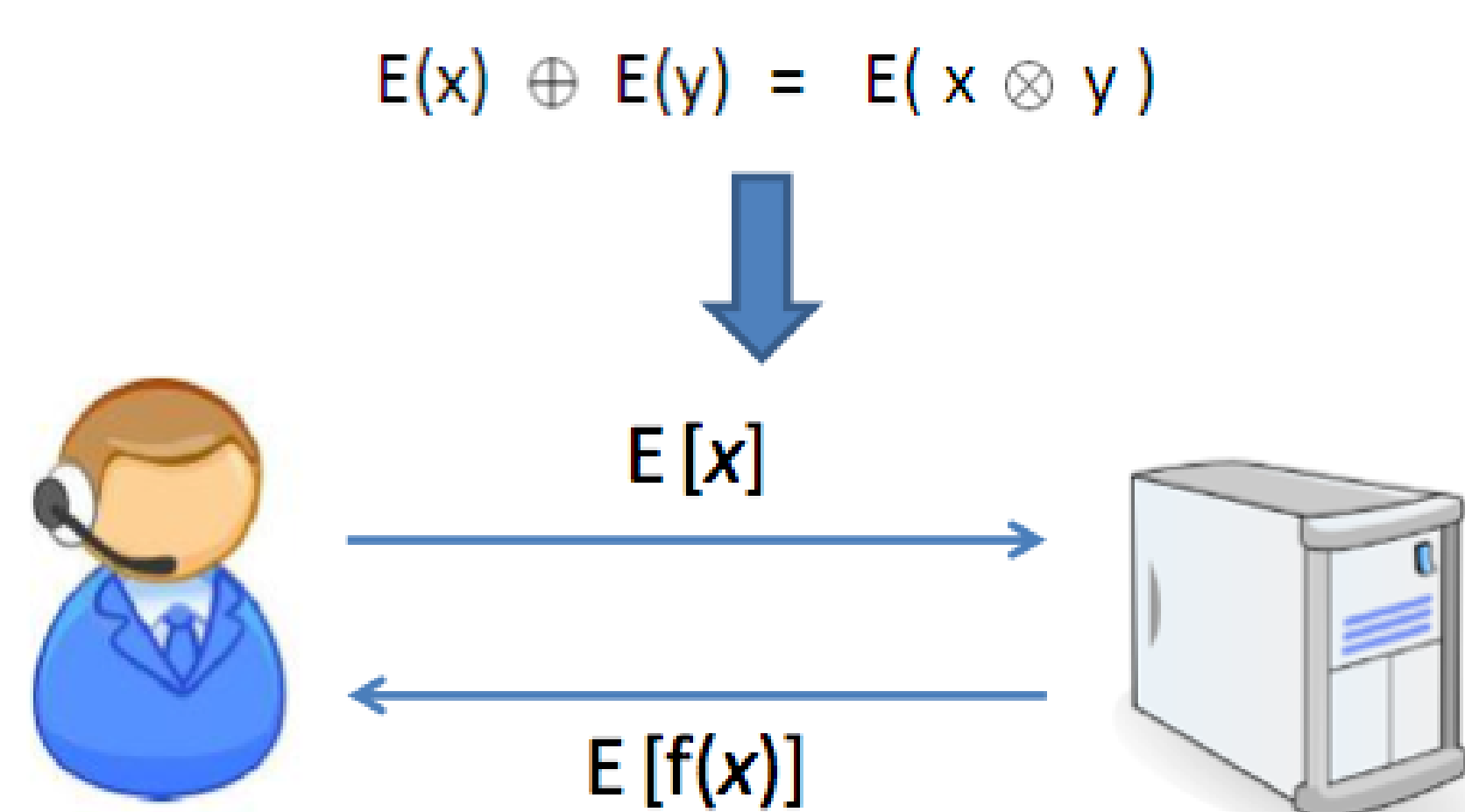
1 Motivation

- It is generally expected that intelligent devices will respond to voice. The voice will often not be processed locally, but relegated to a remote server, as data owners may not have the resources to process their own data. This poses serious privacy risks to the user.
- A person's voice is a legally-accepted biometric, and carries information about their identity, gender, nationality, health, emotional state and a variety of other factors.
- The remote voice service could potentially make undesired inferences about any of these factors, which may be unrelated to the actual service provided.
- This poster discuss "privacy preserving" computational approaches for voice processing that prevent such undesired inferences through cleverly-designed cryptographic and hashing schemes.

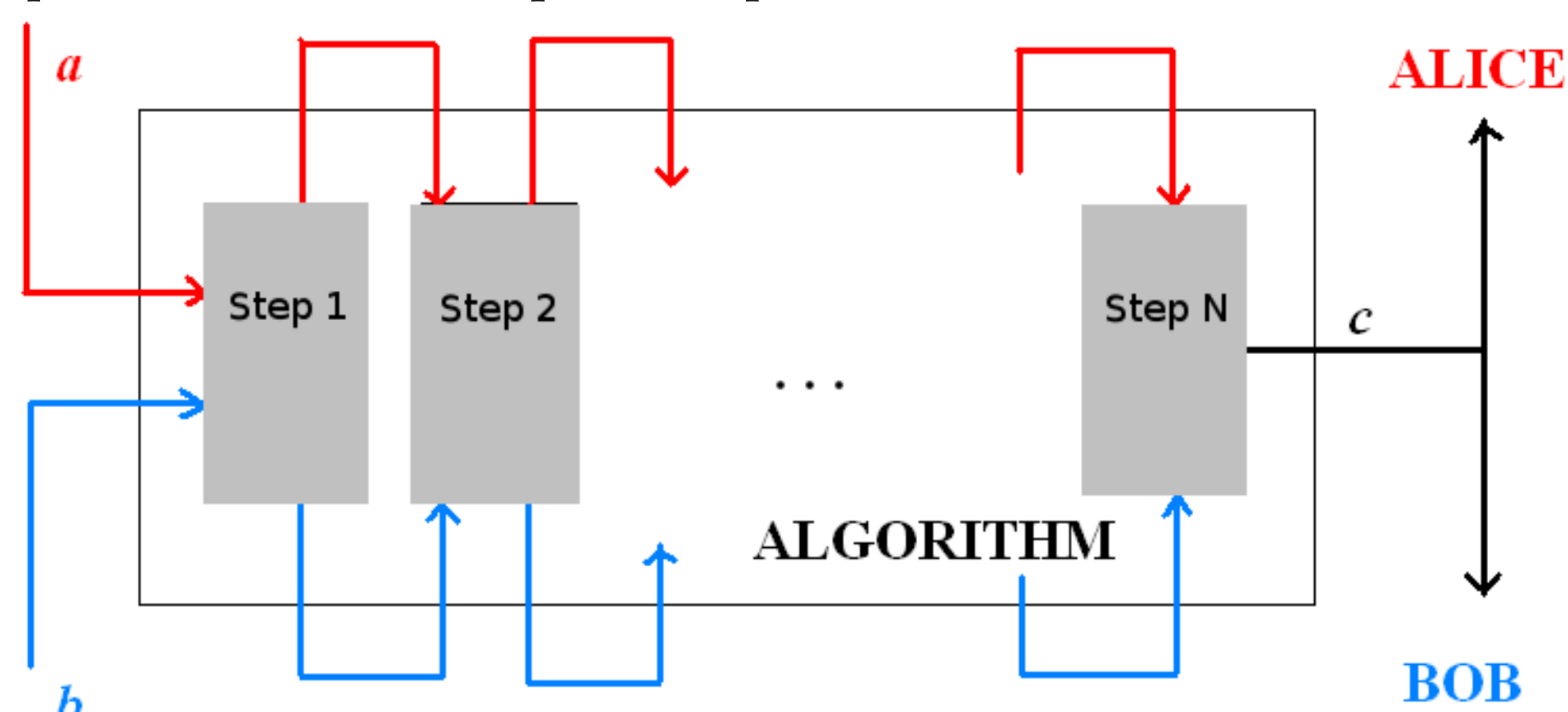
2 Can Cryptography Help?



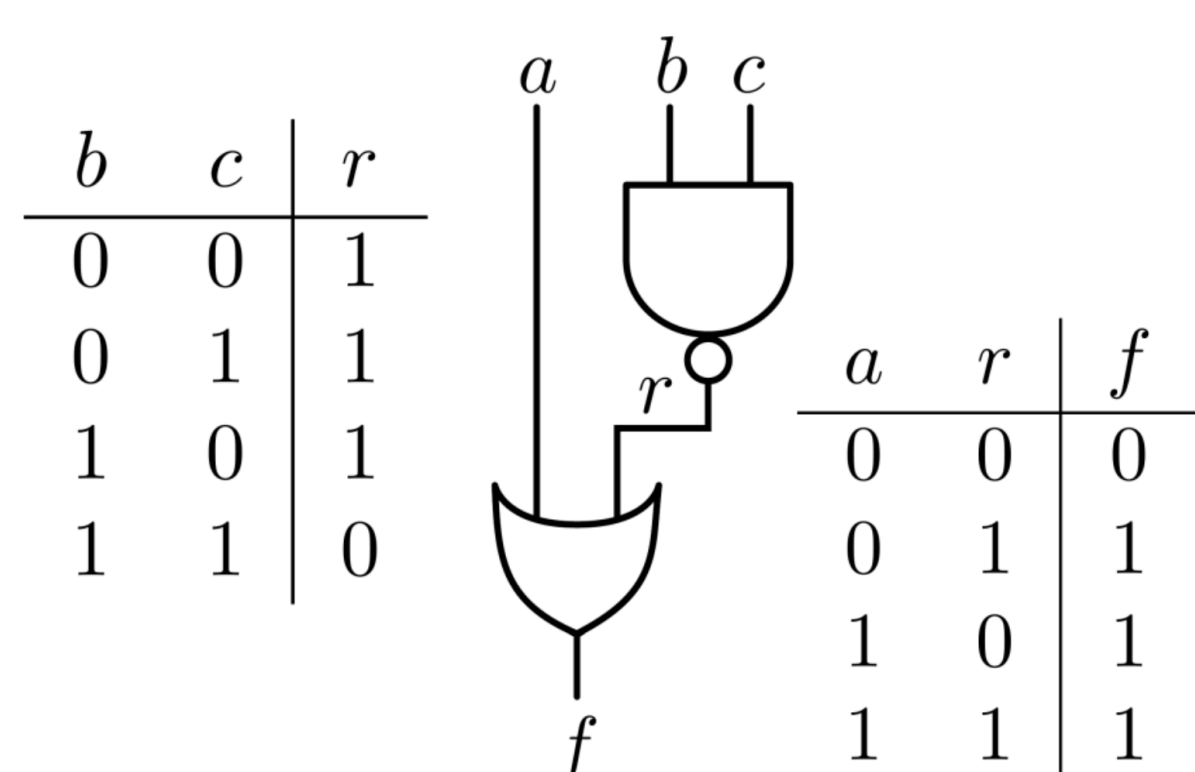
- Secure Multiparty Computation
 - Homomorphic Encryption



– Computation recast as a sequence of primitives



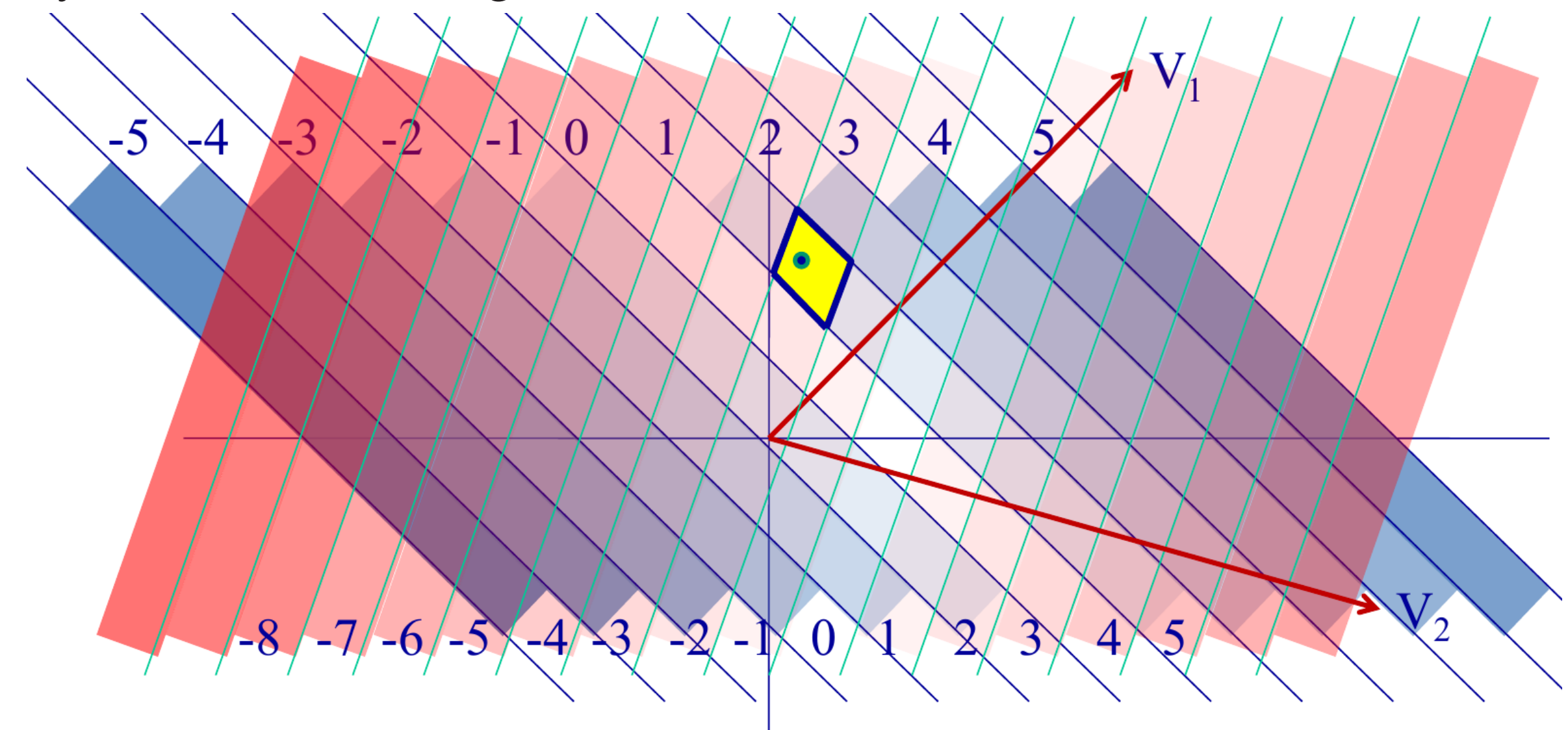
– Garbled Circuits



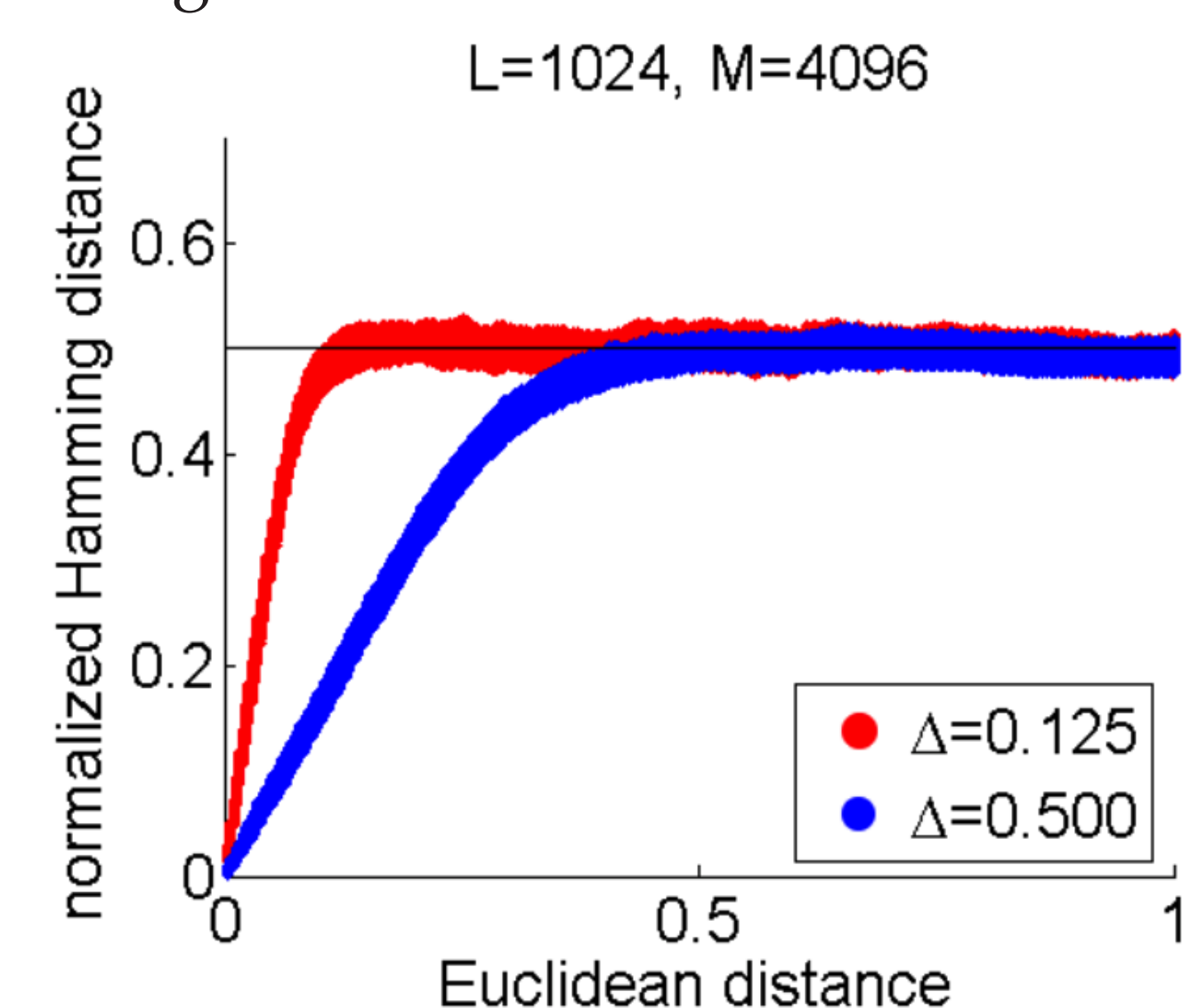
b	c	r	a	r	f
K_b^0	K_c^0	$\mathcal{E}_{K_b^0}(\mathcal{E}_{K_c^0}(K_r^1))$	K_a^0	K_r^0	$\mathcal{E}_{K_a^0}(\mathcal{E}_{K_r^0}(0))$
K_b^0	K_c^1	$\mathcal{E}_{K_b^0}(\mathcal{E}_{K_c^1}(K_r^1))$	K_a^0	K_r^1	$\mathcal{E}_{K_a^0}(\mathcal{E}_{K_r^1}(1))$
K_b^1	K_c^0	$\mathcal{E}_{K_b^1}(\mathcal{E}_{K_c^0}(K_r^1))$	K_a^1	K_r^0	$\mathcal{E}_{K_a^1}(\mathcal{E}_{K_r^0}(1))$
K_b^1	K_c^1	$\mathcal{E}_{K_b^1}(\mathcal{E}_{K_c^1}(K_r^0))$	K_a^1	K_r^1	$\mathcal{E}_{K_a^1}(\mathcal{E}_{K_r^1}(1))$

- Hashing techniques

– Locality-Sensitive Hashing

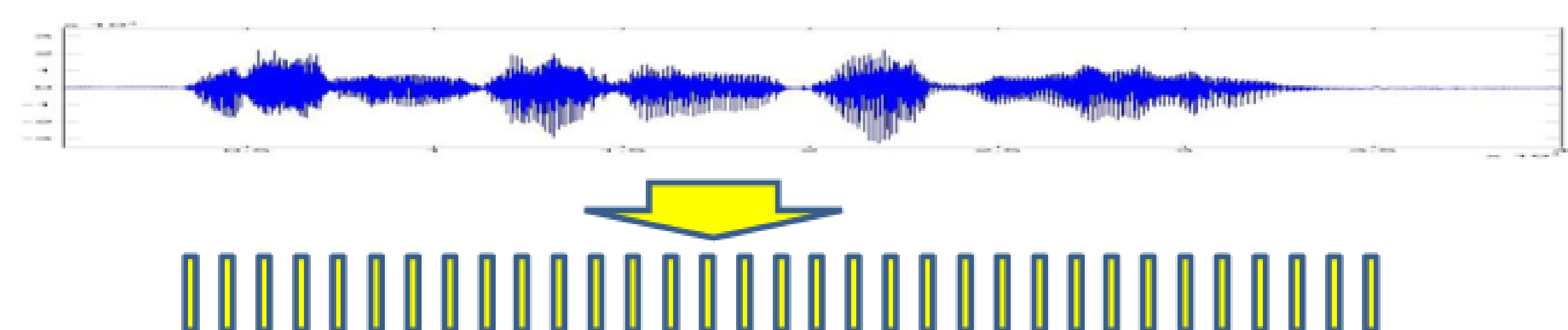


– Secure Binary Embeddings

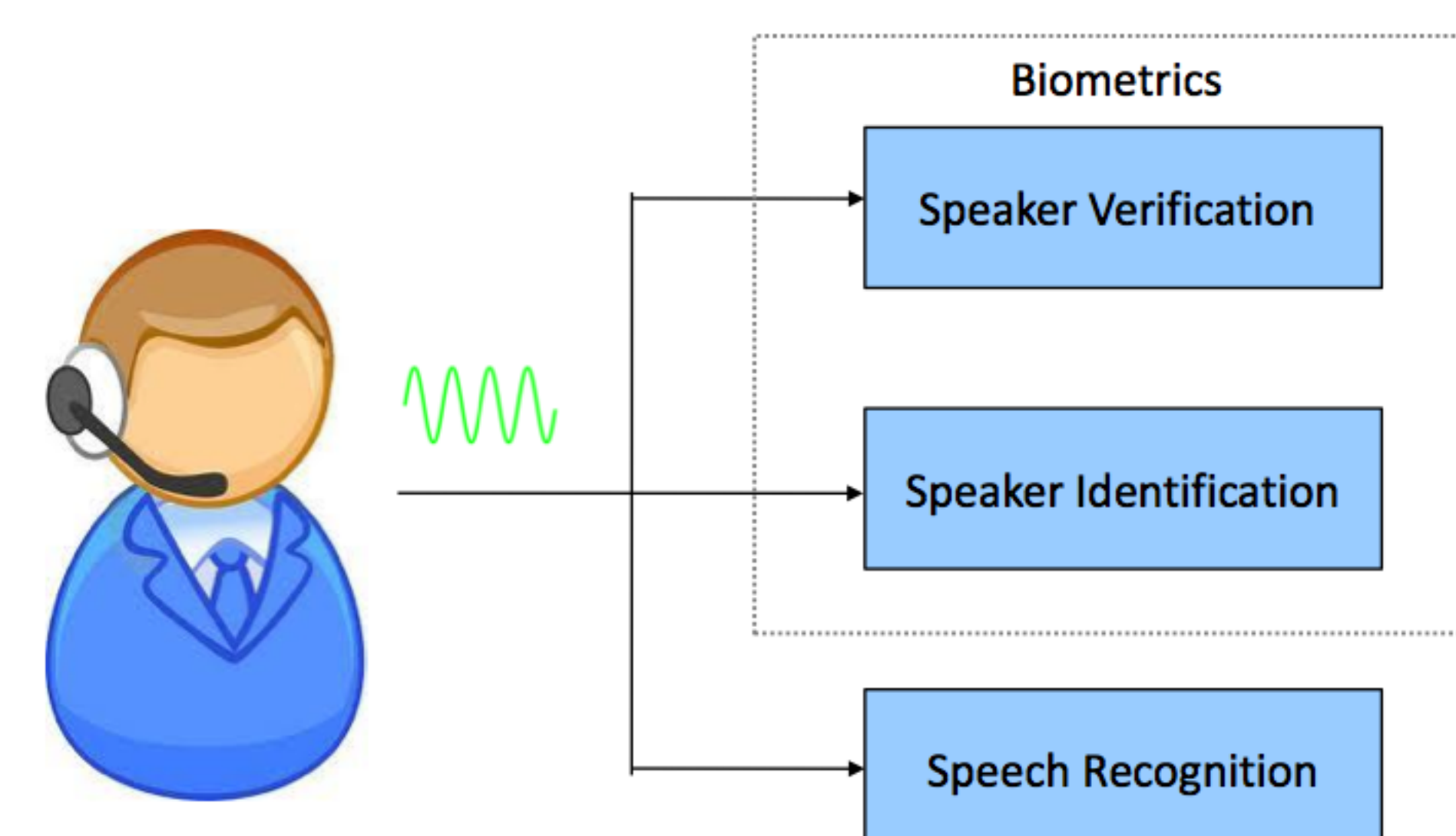


3 Privacy-Preserving Speech Processing

- Work on sequence of feature vectors computed from speech



- Speech processing tasks:



- Speaker Verification: **Are you really** Alice? Yes/No
- Speaker Identification: **Which one of** Alice, Bob, Carol, Dave, ... are you?
- Speech Recognition: **You said** "Hello, world"
- Keyword Spotting: **You said** "blah blah blah ... **drugs** ... blah blah blah"

- Examples:

- Telephone company unwilling to expose audio to intelligence agency
 - * May provide **encrypted** data to the agency
- Agency cannot expose what it is trying to find (a voice, a key phrase) to the telephone company
 - * May provide it in **encrypted** form to the telephone company

- Compromise between obtained results and computational efficiency

Acknowledgments

This work was partially supported by FCT project SUSPECT.