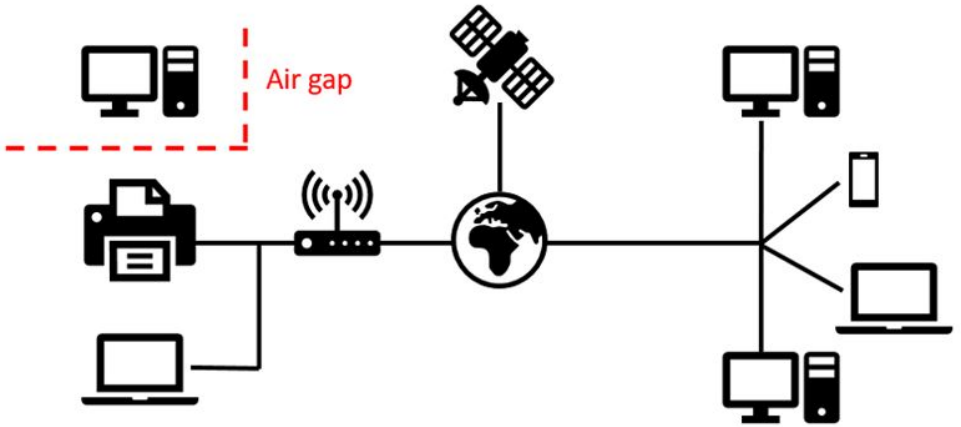
SPyAudio: A Configurable Covert Audio Channel

Arjun Rawal

CMSC 33250

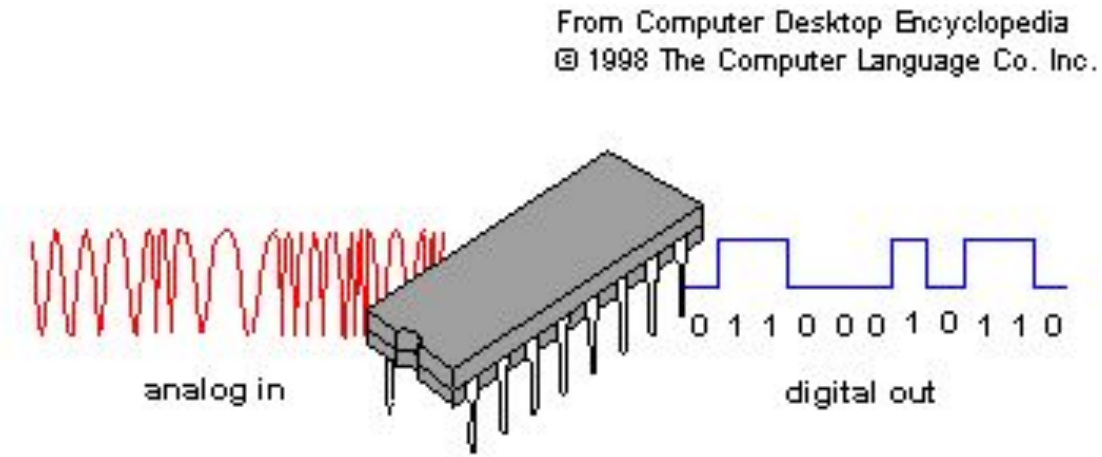
Air-gapped Systems

- Air gapping is the gold standard for preventing unauthorized access
- Demonstrated capacity for attacks through side channels, human error (Stuxnet)
- Can lead to false sense of security



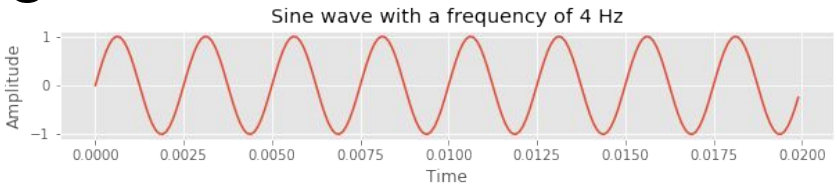
Covert Audio Channels

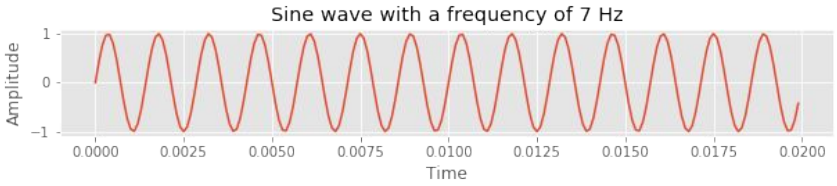
- Basic idea: data can be encoded in audio signal using pitch analysis
- Even if typical transfer methods (Ethernet, Wifi, USB) disabled, audio is often allowed
- Can be done with normal hardware

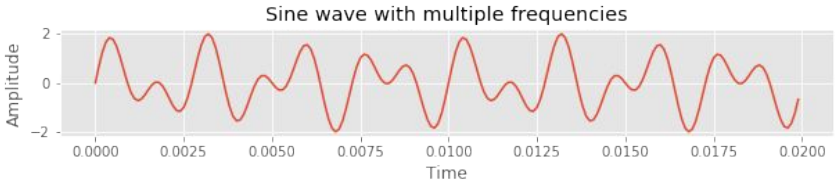


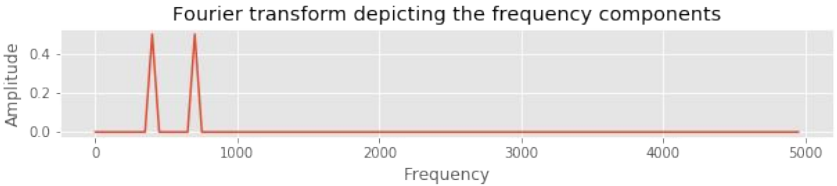
Signal Processing Challenges

- High level of noise and unreliability in audio
- Have to sync up sending and receiving to match phases
- All sound > 18 kHz

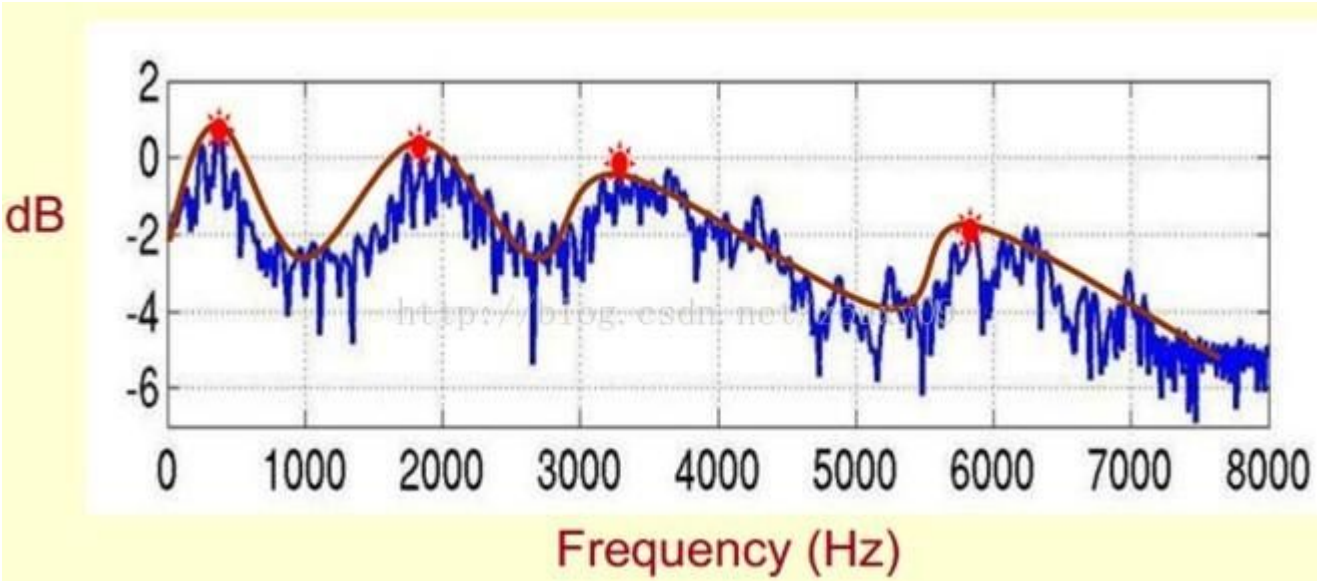








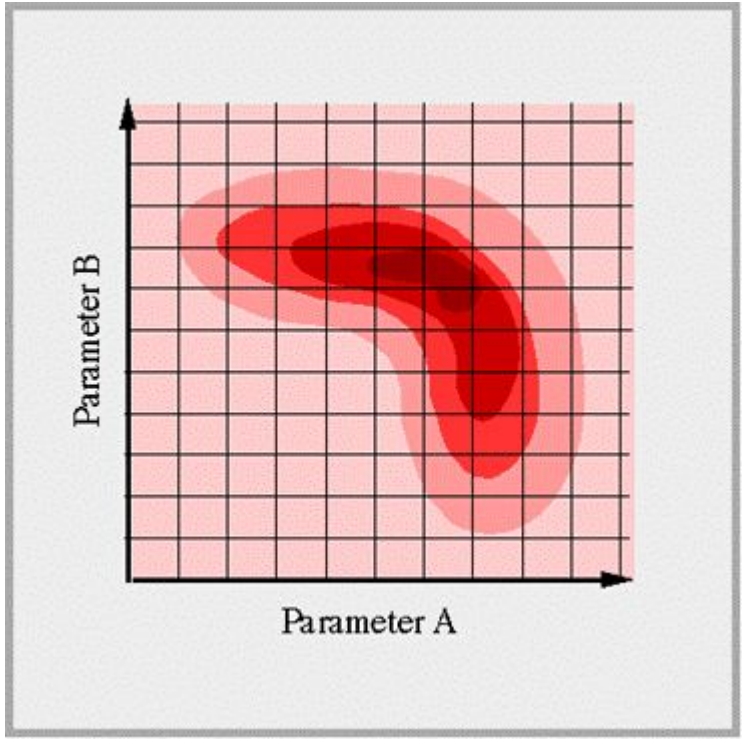
Detecting Signals from Waveform



Benefits of a Configurable Approach

- Other approaches demonstrate in only one environment

- Variety of configurable parameters
  - Base Frequency
  - Frequency Interval
  - Number of Parallel Tones
  - Rate of Tones
  - Error Correction

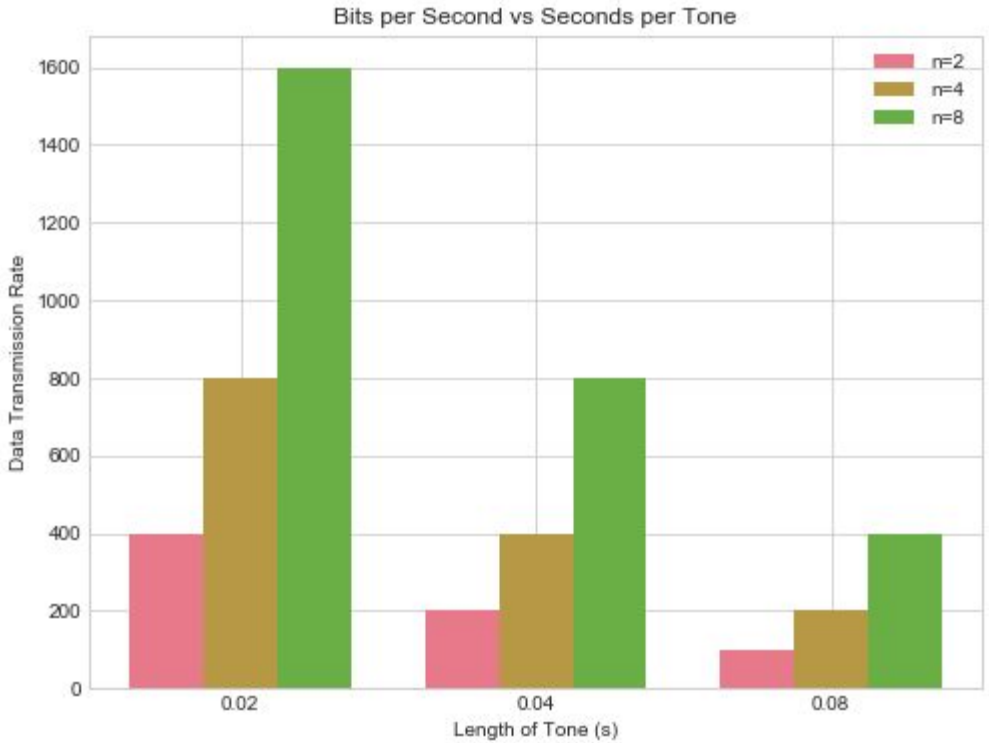


Experimental Methodology

- Commercial technology
  - iPhone 7
  - MacBook Pro
- Framework built in Python
  - numpy
  - PyAudio
  - Wav files
- Used best accuracy of 3 trials

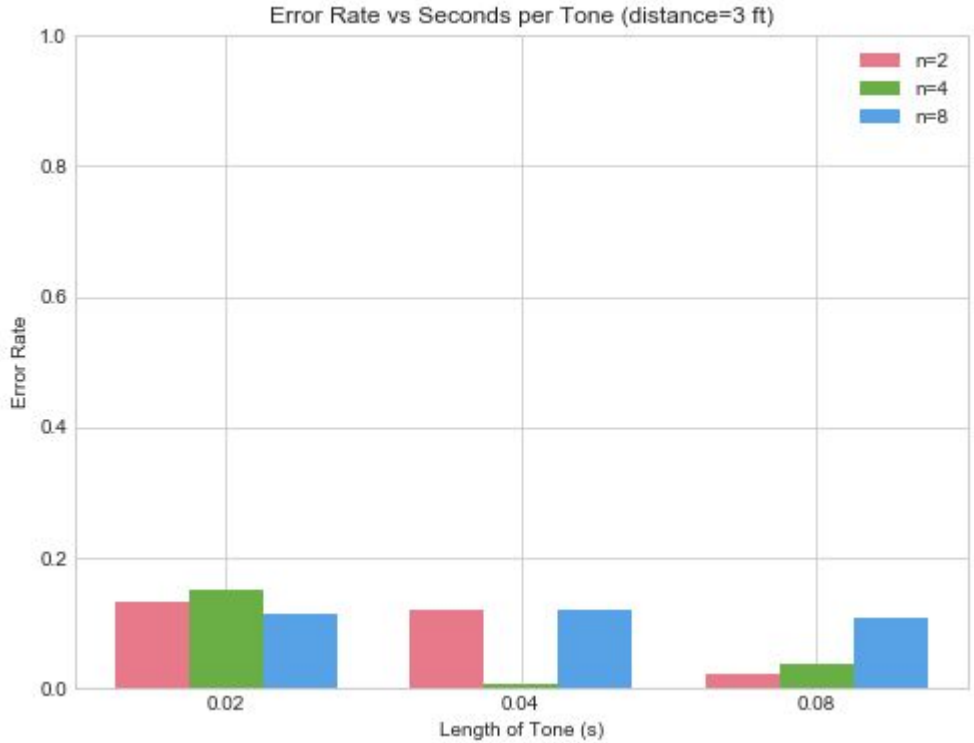
Data Transfer Speed

- Potential transfer speed proportional to length of tone and number of tones
- Non-optimized code capable of significant lossy data exfiltration
  - 1.2 MB/minute
  - Lossy transfer acceptable in some cases



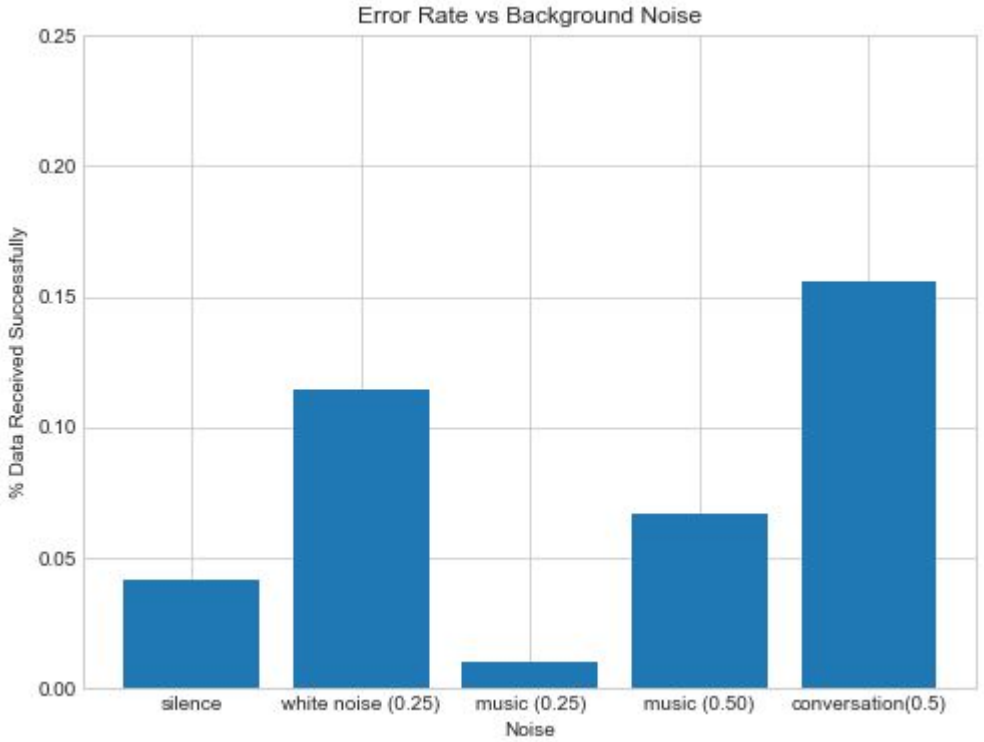
Data Reliability

- General trends
  - More distance => higher error
  - Shorter tones => higher error
  - Parallel tones => ?
- Need to use longer tones at farther distances to maintain reliability
- Max out at 20 feet



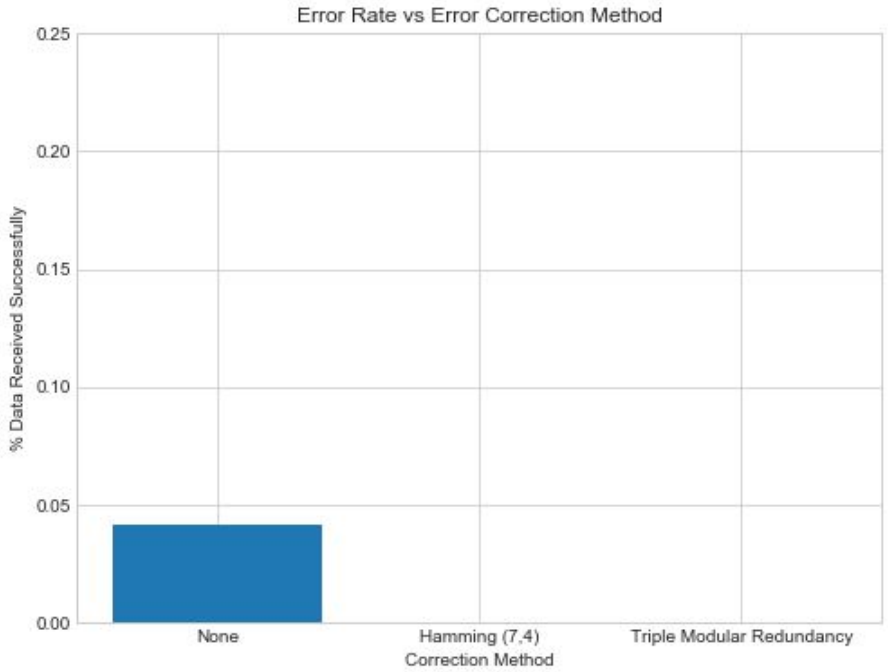
Background Noise

- Somewhat small effect on overall error rate
- Can adapt to background noise by changing tone, adding ECC
- > 18 kHz prevents interference with environment



Error Correction

- Triple Modular Redundancy and Hamming (7,4) both prevent error when used (r=0.04, n=4, d=1)
- Reduces max throughput by 40-70%
- Can adjust to desired levels of reliability

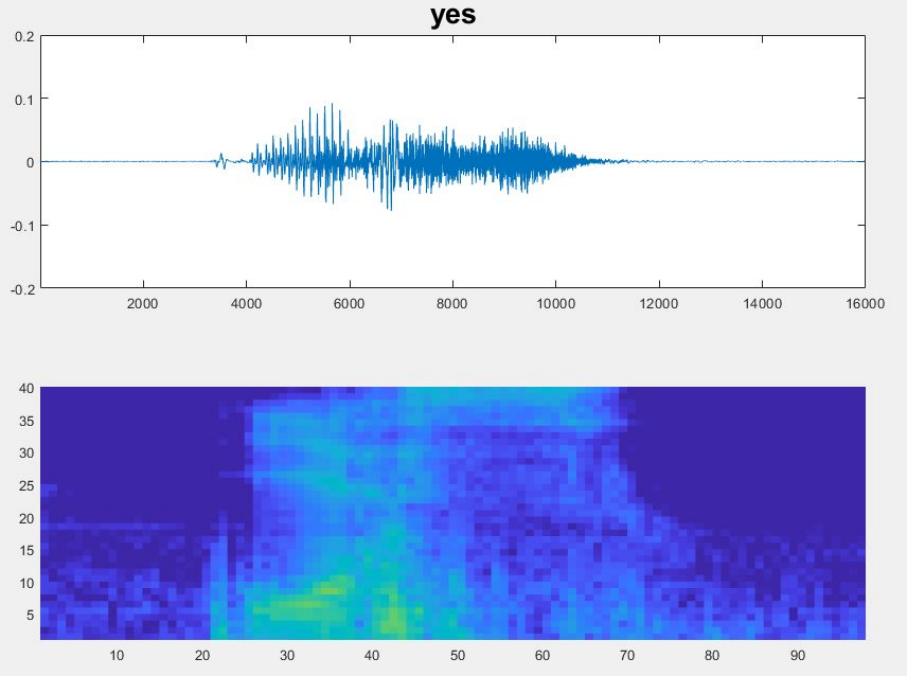


Related Work

- Guri et al. (2018) uses sound to transmit data across a covert audio channel, including from devices that only have speakers
- Novak et al. (2018) focused mainly on transmitting data using sound over very close devices
- Shwartz et al. (2017) covert audio channel between VM and host machine

Future Work: Machine Learning Approach

- Configurable parameters => auto configuration
- Synchronous parameter sweep?
- Auto adjusting stream?
- Use deep learning to detect frequencies with more accuracy than FFT



Questions?