Adaptive Frequency Hopping in BLE

Project Authors Name: Hayden Fuller, Anthony Frias
Project Evaluator Name: Dan Fiumara, Hisen Zhang

- Guidelines to review the report for Question 1-3 below: https://dl.acm.org/journal/dgov/reviewer-guidelines
- Guidelines to review the code artifacts for Question 4-8 below: https://conferences.sigcomm.org/sigcomm/2022/cf-artifacts.html

1. Summary

Provide a brief summary of the project in your own words.

This project investigates and benchmarks channel classification methods for adaptive frequency hopping in Bluetooth Low Energy (BLE) to minimize the impact of heavy interference caused by the crowded 2.4GHz ISM band. Since the BLE specification defines how to hop channels but leaves the algorithm for determining channel quality implementation-specific, innovations can improve channel selection and thus enhance the overall fidelity of the BLE link.

In order to benchmark various channel selection methods, the team implemented a simulation environment, then deployed and characterized five algorithms using it. The results demonstrate that while packet-based methods struggle with responsiveness, classifiers utilizing energy measurements provide earlier interference detection and significantly reduce tail latency. The experiments found the hybrid classifier which accounts for both packet error rate and energy data via an exponentially weighted moving average yields the highest performance. This approach nearly eliminated packet retransmissions and minimized latency.

2. Strengths

Provide strengths or positive aspects of the project.

This project has many strengths.

- Detailed comparison and analysis of methods
    - The project provides a systematic evaluation of five different channel classification algorithms enabling us to draw conclusions about which approach is best and why
    - The approaches implemented also span a spectrum of complexity and aggressiveness which allows us to better analyze tradeoffs between responsiveness and stability under varying interference conditions
- Implementation of reproducible simulation environment

    - The simulation environment implemented in this work enabled fair comparison across approaches, and would make extending this work or testing other algorithms against those explored here much more accessible
    - The simulator is deterministic and uses fixed seeds so that each method sees the same channel, by doing this it remains robust to scientific decision making regarding the efficacy of each approach
- Successfully demonstrates superior performance of hybrid classifier
    - The project clearly shows that the hybrid classifier achieves highest throughput, lowest latency, and least retransmissions as compared to other methods

3. Weakness

Provide any weakness or aspects that can be further improved.

While the project is very strong, there are minor nitpicks that might improve the overall impact of this work.

- Lack of hardware validation
    - The project relies entirely on simulation, ultimately we are deploying these techniques in real systems. To make any strong claims about the true efficacy of any approach studied it must be characterized in hardware
    - This limitation is acknowledged in the report (authors have strong understanding and awareness of this weakness)
- Idealized simulation assumptions
    - The simulation assumes perfect channel state reporting and instantaneous channel map enforcement (also acknowledged by the authors)
- Risk of aggressive channel exclusion
    - The combined classifier was extremely aggressive, this means that it might not be robust to varying interference (also acknowledged)
- Single connection bias
    - In this work only a single connection is managed by the channel classifier. In real world systems a device may maintain connections with multiple devices simultaneously (phone -> earbuds, smart watch, etc.). This may affect the efficacy of the AFH algorithm as the device now needs to update and process multiple independent channel maps while still meeting restrictive timing constraints which could degrade performance
- Lack of sparse traffic analysis
    - Many BLE applications are low-duty sensors which would transmit data very infrequently. This work does not test how these methods perform in sparse data regimes. This is a potentially catastrophic weakness of EWMA based methods which rely on accumulating a history of observations. In sparse traffic the time to

gather sufficient samples to update the EWMA would likely lead to severe lag with the channel hopping reacting far too late to have avoided interference spikes

- Misc.
    - Predictive algorithms (5 & 6) are coded but marked "currently unused" - never evaluated despite being discussed in design section\
    - Algorithm 4 packet loss (4.8% transmission failures) not discussed
    - Success rate vs. throughput tradeoff not analyzed
    - Energy consumption not measured (despite being mentioned as important)
    - Channel starvation scenarios not tested
    - Custom selector (Algorithm 3) is identical to Algorithm 1 - no actual alternative implementation
    - No way to easily run different algorithms - requires manual code editing

4. Documentation: Is the artifact/code sufficiently documented?

Rate from 0% to 100%, where 0% means "documentation is completely insufficient" and 100% means "documentation is absolutely sufficient". If you need to assess both a dataset and tools, please take the average and comment below. In assessing tools, please consider if they are easy or difficult to install/set up and get to run. In assessing datasets, please consider if the meta data is sufficient.

Choices are:

- **1.** 0%
- **2.** 20%
- **3.** 40%
- **4.** 60%
- **5.** 80%
- **6.** 100%

Documentation: Comment on/explain your choice above:

The report is extremely clear. There is only one code file to run and it contains everything necessary to reproduce results. The README is detailed and answers all questions necessary to get the setup running.

5. Completeness: Do the submitted artifacts/code include all of the key components described in the report?

Rate from 0% to 100%, where 0% means "does not include any key components" and 100% means "includes all key components".

Choices are:

- **1.** 0%
- **2.** 20%
- **3.** 40%
- **4.** 60%
- **5.** 80%
- **6.** 100%

Completeness: Comment on/explain your choice above

The single matlab file contains the implementation of the simulation environment as well as all five of the studied channel classification algorithms. It also implemented two predictive algorithms. The implementations appear to be correct and robust.

6. Exercisability: Do the submitted artifacts/code include the scripts and data needed to run the experiments described in the paper, and can the software be successfully executed?

Rate from 0% to 100%, where 0% means "the scripts/software cannot be successfully executed and/or no data is included" and 100% means "the artifact includes all necessary scripts/software and data, and scripts/software (if present) can be successfully executed".

Choices are:

- **1.** 0%
- **2.** 20%
- **3.** 40%
- **4.** 60%
- **5.** 80%
- **6.** 100%

Exercisability: Comment on/explain your choice above

The included matlab file is a standalone and runs with no issues.

7. Results attainable: Does the artifact/code make it possible, with reasonable effort, to obtain the key results from the artifact/code?

Rate from 0% to 100%, where 0% means "no results can be obtained" and 100% means "all results can be obtained".

Choices are:

- **1.** 0%
- **2.** 20%
- **3.** 40%
- **4.** 60%
- **5.** 80%
- **6.** 100%

Results attainable: Comment on/explain your choice above

All results can be easily reproduced. The included file automatically produces all four plots for the selected channel classifier selection. The reproduced result matches the data presented in the paper and supports the authors' claims.

8. Results completeness: How many key results of the paper/report is the provided code meant to support?

Rate from 0% to 100%, where 0% means "the artifact is meant to support no key results" and 100% means "the artifact is meant to support all key results".

Choices are:

- **1.** 0%
- **2.** 20%
- **3.** 40%
- **4.** 60%
- **5.** 80%
- **6.** 100%

Results completeness: Comment on/explain your choice above

The entirety of the results in the report are directly produced by the code.

Reviewer Dan Fiumara, *Dan Fiumara*
Reviewer Hisen Zhang, *Hisen Z.*