PEPPER: Precise Privacy-Preserving Contact Tracing with Cheap, BLE/UWB Capable Tokens

François-Xavier Molina*, Vincent Roca*, Roudy Dagher*, Emmanuel Baccelli†
Nathalie Mitton*, Antoine Boutet‡ and Mathieu Cunche*†

*Inria, France
†Freie Universität Berlin, Germany
‡INSA Lyon, France

Abstract—Contact Tracing (CT) is an old, recognized epidemiological tool, and since a digital variant is now within reach, a variety of smartphone-based solutions have been rapidly developed and deployed since 2020, with mixed results and amid controversies. Yet, achieving reliable and effective digital CT at large scale is still an open problem. In this work, we contribute with an open source software platform on top of which various CT solutions can be quickly developed and tested. More specifically, we design PEPPER, which jointly leverages Bluetooth Low Energy (BLE) and Ultra Wide Band (UWB) radios for contact detection, combined with the DESIRE privacy-preserving CT protocol. We show that PEPPER+DESIRE can operate on cheap physical tokens based on low-power microcontrollers, opening new use-cases with less personal, potentially disposable devices, that could be more widely used. We also evaluate the complementarity of Bluetooth and UWB in this context, via experiments mimicking various scenarios relevant for CT. Compared to BLE-only CT, we show that UWB can decrease false negatives (e.g., in presence of human body occlusion), meaning that more actual contacts will be found, a key benefit from an epidemiological viewpoint. Our results suggest that, while PEPPER+DESIRE improves precision over state-of-the-art, further research is required to harness UWB-BLE synergy for CT in practice. To this end, our open source platform (which can run on an open-access testbed) provides a useful playground for the research community.

Index Terms—Privacy preserving contact tracing, embedded token, Bluetooth Low Energy (BLE), Ultra-Wide Band (UWB)

I. INTRODUCTION

Manual Contact Tracing (CT) is an old, recognized epidemiological tool. Today, the large deployment of smartphones makes a digital variant of manual CT within reach, and solutions leveraging dedicated smartphone apps and Bluetooth Low Energy (BLE) radio scanning have been quickly designed and massively deployed to combat COVID-19 and used in various public and private environments. Such solutions include predominantly the GAEN [1], [2] (inspired by DP3T [3] and implemented in various national applications) and ROBERT [4] (implemented in the French TousAntiCovid application [5]) protocols. However, controversies have flared concerning the downsides of these solutions. Unresolved debates linger regarding their (centralized versus decentralized) design, impacts on privacy [6]–[9], and their general performance/precision [10], [11]. As a consequence, a wide variety of alternatives are being explored.

Motivations for token-based contact tracing — The penetration rate of a CT solution is essential to its effectiveness: if impacts are non negligible even at medium uptakes (e.g., in UK end of 2020, an average 30% app uptake was estimated to avert approximately 1 infection for every 4 infections [11]), the efficiency grows as a quadratic law [12]. Privacy concerns, the availability of a compatible smartphone, and practical impacts (e.g., on battery lifetime), are well identified hindrances to the use of a smartphone for CT.

In order to go further, proposing dedicated physical tokens in addition to a smartphone application could be effective to improve CT uptake, by introducing four major benefits:

• a token is cheap and potentially disposable (e.g., a token can be used for a given event and be reset/recycled);
• a disposable token, per se, is less personal than a smartphone, and thereby raises fewer privacy concerns;
• a well-designed, specialized token is easier to use;
• unlike iOS/Android smartphones, developers can claim full control and transparency over token software and hardware, making fully open, auditable and updatable CT solutions possible for improved efficiency and trust.

However a token raises additional challenges in terms of energy, processing, and communication, and asks for appropriate design tradeoffs. It is important to note that not all CT protocols are compatible with the resource constraints of tokens. For instance, a fully-decentralized risk analysis solution such as GAEN, that requires each device to daily download Megabytes of keys identifying COVID-positive users and perform intensive computations for pseudonym matching is not adequate for low-power tokens (see Section II).

Furthermore, privacy is a key requirement for a large scale adoption, and from this viewpoint the present work largely differs from token-based initiatives that either target workplaces or require a non-anonymous user registration (see Section II). Last but not least, studies show that BLE, as used in this context, tends to not provide solid and reliable distance estimations [13] or [14], which pave the way to combine different technologies to improve the reliability.

This research has been partially supported by the ANR17-CE25-0014 CISC project (https://anr.fr/Project-ANR-17-CE25-0014) and by the ANR-20-CYAL-0002 PIVOT project (https://anr.fr/Project-ANR-20-CYAL-0002). The authors are grateful to the members of the RIOT-fp project financed by Inria, that aims at developing a cybersecurity-focused distribution of the RIOT operating system (https://future-proof-iot.github.io/RIOT-fp/).
Contributions of this work — The present work explores three areas. A first area of work is enabling the use of tokens and the associated design tradeoffs. A second area of work concerns the adaptation of privacy-by-design CT to token-based hardware, software, and network constraints. A third area of work consists in complementing BLE radio scanning with Ultra-Wide Band (UWB) proximity detection [25].

Based on our observations, we pick a state-of-the-art privacy-preserving CT protocol using BLE, DESIRE, and:

- we show and analyze how DESIRE can be adapted to accommodate resource-constrained embedded devices such as cheap tokens based on low-power microcontrollers, and we provide the first open-source implementation of DESIRE suited to these environments.
- we design and implement PEPPER, a technique to efficiently synchronize UWB Two-Way-Ranging (TWR) exchanges by exploiting BLE advertisement and SCAN events. PEPPER remains privacy-compatible, requires little-to-no overhead in the BLE advertisement payload, and requires only a small 2ms UWB receive (RX) window on the responder side.
- with our PEPPER+DESIRE Proof-of-Concept (PoC), we demonstrate how a state-of-the-art privacy-preserving CT protocol based on BLE can be enhanced with a complementary, backward-compatible, mechanism providing accurate UWB Time-of-Flight distance estimations.
- we evaluate PEPPER+DESIRE experimentally in several scenarios relevant for CT. In particular, UWB-based classification can decrease False Negatives (e.g., in presence of occlusion due to human activity), meaning that fewer actual contacts are missed compared to BLE, which is a nice result from an epidemiological viewpoint, although it comes with higher false positives (the "price to pay").
- to this end, we provide an open source software platform on which novel CT solution can be quickly developed, leveraging jointly BLE and UWB radios on low-power microcontrollers, and which can be deployed remotely on an open-access IoT testbed for experimental research.

Non goals of this work — This work does not try to assess whether the use of physical tokens for CT is more acceptable than an application through user studies (out-of-scope).

It does not try to assess whether CT can mitigate COVID19 (it is left to epidemiologists, and [11] concluded that CT saved 4200 - 8700 lives in the UK in September-December 2020).

It does not try to assess whether measuring distances more precisely improves the exposure risk assessment.

It does not consider BLE or UWB security (e.g., in front of an attacker who lies during UWB TWR exchanges to artificially increase or reduce distance measurements).

Paper outline — After discussing related works in Section II, we give background on CT with DESIRE and BLE in Section III. We then detail the design of a UWB extension for DESIRE in Section IV. We provide insights of the underlying open platform and the PEPPER+DESIRE PoC in Section V and Section VI respectively. We evaluate the performance of our PoC via experiments in Section VII. We finally discuss results in Section VIII before concluding.

II. RELATED WORKS

The dominant CT protocols (e.g., GAEN [1, 2] and ROBERT [4]) and their implementations target the user’s smartphone, using BLE radios. Fewer works target low-power hardware such as microcontroller-based tokens, or alternative radios such as UWB.

Token-based contact tracing — Several token-based CT initiatives have been developed and commercialized for use in work-places. A number of products such as SaferMe [1] or Estimot [2] propose token-based solutions for entreprises. These solutions use BLE and proprietary, adhoc contact protocols with minimal privacy considerations. However such approaches are incompatible with large scale deployments and strict data protection regulations (e.g., GDPR).

A nation-wide system, TraceTogether, was deployed in Singapore both as an Android/iOS application and autonomous token [1] When a person is tested COVID+ their contact history stored on the device is uploaded to the server, their contacts identified and called (manually). Here also the CT protocol features minimal privacy considerations (the authority directly knows who met who).

Call for proposals were launched at the European level to develop token-based solutions compatible with GAEN system [3]. However the problem is complex due to the decentralized approach for risk-analysis of GAEN. In particular, a bottleneck appears whereby tokens need to continuously download and process long lists of keys and pseudonyms of COVID+ users (several MBytes per day for peak infection rates). Not surprisingly, we are not aware of any product or even proof-of-concept resulting from this initiative.

These related works highlight (1) the difficulty of designing embedded tokens for CT with high privacy expectations, compatible with strict data protection regulation, and (2) the need for a flexible CT protocol that is suited to both token and smartphone environments.

Proximity detection using BLE & UWB — Prior studies have shown that BLE cannot provide accurate distance measurements. Indeed, BLE has not been designed for this goal, its accuracy is highly dependent on the device location, orientation, advertisement channel and, for the receiver, on the knowledge of the transmission power, as analyzed in studies such as [14, 16], which also highlight that a careful calibration step is paramount. Even with good calibration, field
results have shown this to be insufficient to precisely identify contacts \cite{10, 13}. The main advantages brought by BLE are that it is relatively cheap in terms of energy consumption, it is widely available (e.g., on smartphones), and semi-passive overhearing RSSI-based approaches are possible (broadcast transmissions), making it to scale as a factor of $N$ with the number of nodes in the vicinity.

On the opposite, Ultra-Wide-Band (UWB) enables accurate distance measurements where a device can accurately measure the Time of Flight (ToF) with a nearby target device. Using relatively cheap commodity UWB hardware such as a DWM1001 \cite{17}, the initiating device can derive the distance with an error bounded to a few tens of centimeters. Compared to a BLE-based approach, techniques using UWB perform better in face of multi-path fading effects and interferences, however, UWB (specially HRP UWB \cite{18}) incurs a higher energy consumption. We also note that round-trip, point-to-point interaction required by some UWB approaches achieving better accuracy scales as a factor of $N \ast (N - 1)$ with the number of nodes in the vicinity, making it a challenge for crowded environment. No dual radio approach exists in prior work except Janus \cite{19}, which designs and studies a combination of BLE and UWB aiming to provide a workplace CT solution, with minimal privacy requirements.

Closest related works — To the best of our knowledge, the closest works are DESIRE \cite{20, 21} and Janus \cite{19}. Compared to DESIRE, our work extends the protocol to leverage jointly UWB and BLE, instead of only using BLE, and adapts the protocol to fit on low-power microcontrollers, not only on smartphones. We also provide an implementation and report on experiments. Compared to Janus, while our experiments confirm the bulk of performance assessments the authors report (e.g., on distance estimation accuracy, energy consumption), our work targets nation-wide deployments with much more stringent privacy requirements. For example, our work does not rely on simplifications such as long lasting node identifiers, or non-cryptographic bitmap carried in BLE packets to identify nodes for UWB ranging.

III. BACKGROUND

This section introduces useful terminology and the DESIRE CT protocol at the core of our proof-of-concept.

A. Contact Tracing Protocol Terminology

We use the following definitions:

- **Neighbor**: any device whose advertisements are heard over BLE, regardless of its proximity.
- **Encounter**: a neighbor that fulfills some conditions that would merit collecting proximity information. With DESIRE, at a minimum, it is the ability to reconstruct the "Ephemeral Bluetooth IDentifier" (EBID) of the neighbor, requiring a proximity that is at least 40 seconds long.
- **Contact**: an encounter which, at the end of an "epoch", fulfills a set of conditions (e.g., based on an average measured RSSI or distance, or a duration) that warrant such encounter to be logged. This notion does not necessarily imply a COVID exposure risk, since the remote user will not necessarily be tested COVID+ in the following days, and the local criteria used to tag an encounter as contact are different from that of the risk evaluation process.

- **Proximity Discovery**: process by which the status of a device is progressively refined, from neighbor (detected), to encounter (EBID reconstructed), and potentially contact.

B. DESIRE for BLE

The DESIRE CT protocol \cite{20, 21} fundamentally relies on the concept of "Private Encounter Token" (PET), a cryptographically generated identifier which uniquely identifies an encounter between two devices. Similarly to GAEN or ROBERT, each device periodically broadcasts its device pseudonym, called "Ephemeral Bluetooth IDentifier" (EBID) via BLE, changing at the end of each epoch (by default 15 minutes). However, the only purpose of EBIDs is to enable two devices that discover they are close-by to locally and secretly compute the same PET, using the Diffie-Hellman key exchange protocol. As a consequence, PETs are only private to the two devices who met, leaving passive eavesdroppers powerless. PETs are also unique to each encounter and change across time and devices (e.g., if A and B meet more than 15 minutes apart, different PETs are generated, and if A meets B and C simultaneously, different PETs are generated for each encounter). PETs can therefore be shared with a server to query the user exposure status without compromising privacy. Similarly, if a user is tested COVID+, disclosing her PET history across the past 14 days is sufficient to inform the persons she met of a potential exposure risk\cite{6}

When compared to other deployed CT techniques, such as GAEN or ROBERT, the DESIRE approach is a total paradigm shift in that it does not identify devices but encounters, which significantly improves privacy.

This paradigm shift has additional benefits in terms of flexibility and interoperability. Indeed, assessing an exposure risk is essentially a matter of looking for matches between two lists of PETs: the device contact history across the past 14 days and a "black list" of PETs collected from users tested COVID+ who agreed to share their own PET history. If follows that DESIRE enables both centralized, decentralized or hybrid risk evaluation models, depending on where this pattern match is performed. In any case, deployments following different risk evaluation models are fully interoperable with one another, meaning two different models can be used within the same deployment. We show in section \cite{IV} how this unique flexibility feature benefits to a token-based deployment of DESIRE.

Finally, it should be noted that each EBID is 32 bytes long, too large to fit into a single BLE Advertisement message (16 bytes payload). Therefore, an EBID is split into three slices and BLE messages (a slice identifier is needed, hence the third one), each message being transmitted in a dedicated carousel round during 20 seconds, as depicted in Figure \cite{1}.

\footnote{For privacy reasons, two PETs are generated per encounter, one to query the user status, another one to report a positive COVID+ status \cite{20, 21}.}
Reconstructing an EBID requires collecting at least one BLE message of each round, taking a minimum of 40 seconds (if a message of the third round is immediately received). This is not an issue from a risk analysis viewpoint, as a contact requires a minimum of 5 to 15 minutes proximity, depending on local epidemiological criteria.

IV. DESIRE on UWB Capable Tokens

This section explains how DESIRE was adapted to the case of UWB capable embedded tokens, while keeping a full interoperability with a smartphone CT application. It first discusses design considerations, then continues with several protocol considerations, going further into details.

A. Design Considerations

Several key considerations framed the design:

- **Storage, computation and communication capabilities are fundamentally limited**: These considerations impact how the DESIRE CT protocol is deployed, where risk evaluation is performed, and the balance between BLE and UWB technologies.

- **BLE remains essential**: BLE is essential for neighbor discovery, in order to perform a first proximity filter, and as a mechanism to schedule UWB ranging rendez-vous if both nodes are UWB capable. Being a broadcast technology (through advertising), it also scales better than UWB (that leverages unicast communications), with lower energy consumption. And last but not least, BLE is essential for backward compatibility with BLE-only devices (e.g., most current smartphones).

- **UWB is a key technology**: one reason is of course its accurate distance evaluation capability inherited from Time of Flight (ToF) measurements. Additionally the distance computation is much easier with UWB data than it is with BLE RSSI measurements, both in terms of computation and data storage (i.e., the BLE RSSI-to-distance model relies on a probabilistic approach requiring expensive floating point arithmetics and an order of magnitude more data to be kept [21]). This is essential for the embedded platforms PEPPER targets.

- **Decoupling protocol mechanisms from policy and heuristic considerations**: when it comes to defining UWB ranging rendez-vous, several heuristics are possible to reduce energy consumption to an appropriate tradeoff, depending on the exact situation.

B. DESIRE Adaptation Details

This section now details how the above design considerations framed the DESIRE adaptation.

1) **Risk Evaluation Versus Deployment Model**: The concept of PET enables several types of architectural deployment variants, a centralized (denoted D1 in [21]) versus decentralized (denoted D3) risk evaluation design, as well as an intermediate stateless approach (denoted D2). Of course, each variant impacts what data (PET) is shared between devices and the infrastructure. In the present work, we show that this flexibility is also a key advantage for an embedded token deployment of a privacy-preserving CT protocol.

A token needs to limit as much as possible processing and communication, or alternatively offload as much of it to a more capable entity (proxy), server or smartphone, that also benefits from an Internet connection (not likely possible on an embedded token without a high energy cost).

In the case of the D1 and D2 architectures, the token would only require to be associated with a proxy

Thanks to this "token-proxy" channel, tokens can (1) periodically upload their encounter data (PETS+ BLE/UWB metadata) thus allowing on-edge risk evaluation, and (2) send exposure status requests via the same channel.

In a fully decentralized approach (D3), risk evaluation and PET matching are costly operations that cannot be performed on the Device. These operations can be offloaded to a trusted Proxy (e.g., the token bearer own computer or cellphone), which is conceptually close to a D2 deployment.

2) **Proximity Discovery**: Let us now focus how UWB and BLE are jointly used for Proximity Discovery. We assume devices A and B are close enough to hear each other via BLE.

The process depicted in Figure 1 reflects how proximity is progressively refined: from neighbor detection, to encounter (EBID reconstruction) and then UWB ranging rendez-vous.

More precisely, A transmits its $EBID_A$ slice per slice, within a carousel (there are several rounds in the carousel, one per slice, i being the round index); round 1 starts at time $t_{1A1}$ and lasts 20 seconds (default), round 2 at $t_{2A2}$, round 3 at $t_{3A3}$, and so on, until the end of the epoch. In this phase, two

![Fig. 1: From neighbor and encounter discovery via BLE, to distance measurement via UWB (high-level view) ](image)
criteria determine if a neighbor is considered an encounter:
- the neighbor EBID must have been reconstructed,
- and the neighbor must remain visible over BLE, i.e., if after successfully reconstructing $EBID_A$ no more advertisements are received for a large enough time, the device is no longer considered an encounter until a new BLE advertisement is received.

These rules privilege passive operations for encounter discovery, and schedule (resp. unschedule) UWB ranging interactions once the encounter is confirmed (resp. lost).

The exact heuristic behind this local decision algorithm is out of the scope of the present work, but it is expected to favorably reduce UWB traffic and battery consumption (we evaluate some of them in section VII).

Fig. 2: From BLE advertisements to UWB rendez-vous

3) Defining UWB Rendez-Vous in PEPPER: When an encounter between A and B has been created (which implies A and B reconstructed their respective EBID), the next step is to deterministically schedule a rendez-vous to perform UWB TWR exchanges. Since a DESIRE BLE advertisement packet already uses all the available space (16 bytes plus a few extra bytes in the trailer) in the payload to transmit the EBID slices, and in order to avoid additional traffic, we leverage the individual advertisement/scan events as synchronization points. This rendez-vous synchronization procedure is central to PEPPER. Figure 2 illustrates it. Let us assume that A is the initiator. Its BLE traffic defines regular time references: $t_{A,n,i}$ that correspond to the end of an advertisement procedure (i.e., the end of transmission of a BLE packet on three channels, see Section VI). These events are depicted as short light-blue pulses in Figure 2.

Every time the BLE scanning window of B (long blue pulse) overlaps with a given advertisement, then A and B share the same $t_{A,n,i}$ time reference. Then, A and B both compute the waiting time to add to this time reference:

$$delta_{AB} = f(EBID_A, EBID_B, slice\_index)$$

and A sends a UWB ranging request (short orange pulses in Figure 2) at times: $t_{A,n,i} + delta_{AB}$. On its side, B opens a RX window (long orange pulse) at roughly the same time, resulting in a successful TWR exchange.

By construction, $delta_{AB}$ depends on numerical values $EBID_A$ and $EBID_B$, so that different devices end up with different delta values even if originating from the same A device. This is important to reduce collisions risks across different pairs of devices.

The protocol does not guaranty that A and B actually agree on starting a mutual UWB ranging at a given rendez-vous; instead it defines accurate time references when such a UWB ranging can happen. If B decides (on its own) not to engage into UWB ranging, it is sufficient for B to ignore UWB traffic (i.e., do not listen to UWB interface) at a given rendez-vous. The situation is similar for A that may not start the UWB message exchange. It can also happen that B fails to scan a given advertisement, and while A attempts to initiate an exchange, B is not present at the rendez-vous (e.g., $t_{A1,4}$ in Figure 2). This is not a major issue as the initiator request is short lived and cheap compared to the responder’s UWB RX window. Discovery is also asymmetric, A or B may attempt to initiate TWR exchanges before its counterpart is ready.

Examples of criteria to decide whether a ranging between A and B is appropriate or not include:
- TWR expiry-time: time before a ToF/Distance value should be refreshed;
- configuration preferences, for instance in order to save battery, like a minimum period between two successful TWR rangings with a remote node;
- a rather small RSSI value that may suggest the remote device is probably too far away (low epidemiological risk, an accurate distance estimation is useless).

The exact heuristic behind this local decision algorithm is out of the scope of the present work, but it is expected to favorably reduce UWB traffic and battery consumption (we evaluate some of them in section VII).

4) UWB Device Identification: Then, the TWR protocol needs to identify each device during the UWB ranging. To that purpose, each node generates an IEEE802.15.4 "short address" from their EBID, as well as a common PAN-ID, by truncating to 16 bits the SHA-256 hash of both EBIDs. The {short address; PAN ID} pair being 32 bits long, the risk of collision can be neglected, especially as the potential collision time span is limited to the TWR transaction.

5) UWB Single-Sided Two Way Ranging: Most common TWR protocols are Single-Sided TWR (SS-TWR) and Double-Sided TWR (DS-TWR). DS-TWR reduces the influence of clock-skew, resulting in more accurate ToF values, but requires an extra message. With SS-TWR, only the initiator can compute a ToF (Time of Flight) value, while with DS-TWR, it’s the responder. For this reason, often, a final "report" message is used to communicate the estimated ToF to the counterpart.

A privacy issue for CT, is that there is no encryption, and therefore any passive eavesdropper can deduce from this message the distance between the ranging parties. The IEEE802.15.4z [23] specifies ways of protecting this data through symmetric cryptography. In our context a symmetric
key could be derived similar to the $PET^1$ and $PET^2$ computations (see [21] for details), by having A and B compute:

$$PET^3 = H(\langle S \rangle | g^{X,Y})$$

Another alternative is to omit the “report” message altogether, in which case only the initiator can estimate the ToF, saving a message per TWR exchange but requiring an extra exchange so that the counterpart also obtains a ToF value. This is the approach used in our proof-of-concept.

6) Contact Registration: CT protocols usually determine relevant encounters based on encounter duration and in some cases RSSI thresholds [3]. But RSSI variability does not allow to accurately categorize contacts, which is why DESIRE uses a probabilistic risk analysis to weight in the different RSSI values [22]. By relying on distance values CT application could use the World Health Organization 2m inter-personal distance as a criteria to filter out non-relevant contacts. Doing so avoids unnecessary PET generation (saving CPU) and reduces storage and upload bandwidth requirements (see Section VII-C).

V. OPEN PLATFORM

We implemented the PEPPER+DESIRE experimental platform with fully open components: (1) Open Hardware, (2) Open Firmware building blocks and (3) Open Testbed for online experiments. The source code of the experiments is shared with the community, along the datasets, and the application source code on a dedicated companion repository [24]. The software we designed is usable on the open access testbed FIT IoT-LAB [25] for experiment repeatability, and thus as an open platform for any future work.

Fig. 3: Decawave DWM1001 module and software stack

1) UWB/BLE Open Hardware & Testbed: The prototype we use is shown in Figure 3. More precisely, the prototype uses a commercially available development kit embedding a small Decawave DW1001 module [26] and an ARM Cortex-M4 microcontroller with 512kB flash memory and 64kB RAM operating at 64MHz clock frequency. We designed a wearable casing embarking a small battery to carry around our prototype for our experiments. The same hardware is deployed in the open-access testbed FIT IoT-LAB [27], where up to 24 such boards (DWM1001-DEV) can be used remotely for experiments. A 3D map indicates the precise positions of these devices in the testbed, which can be used for early developments and scalability tests. The open source embedded software stack we provide (see below) can be flashed remotely and executed on the testbed.

2) Open Source Software Stack: The prototype leverages on a fully open source embedded software stack, based on:

- RIOT [28]: a general-purpose low-power open source OS for microcontroller-based IoT.
- UWB-core [29]: an open source library from Decawave, providing UWB hardware abstraction layers, an UWB media access control (MAC) layer, and ranging functions (RNG): SS-TWR, DS-TWR, N-TWR.
- NimBLE [30]: an open source library from Apache Mynewt, providing a Bluetooth 5.1 stack (both Host & Controller), including advertising extensions and various low level API access.

Our implementation thus provides an entirely customizable software framework with full access to all the embedded hardware functionalities and the network stacks.

VI. PEPPER+DESIRE IMPLEMENTATION

We now detail our PEPPER+DESIRE Proof-of-Concept (PoC) and choices made, specifically regarding Proximity Discovery and its impact in Contact Registration.

Although risk evaluation and device enrollment in the CT system are largely considered out-of-scope, a demonstrator is provided in the companion repository. It illustrates the concept of proxy described in section V.B via CoAP over IPv6/BLE endpoints, for offloading encounter data (namely ETL, EPPER) as well as querying exposure status requests and infection declaration.

**Proximity Discovery:** The neighbor discovery step involves periodic BLE advertisements and scanning that should be balanced to trade-off between latency and energy consumption. For the advertisement profile, we adopt DESIRE’s default parameters: EBID slice is advertised every second, EBID slice rotation every 20 seconds, EBID rotation every 15 minutes. As for the BLE scan profile, we adopt Android’s “Balanced” scan mode that scans approximately 25% of the time.

For simplicity and to avoid CPU intensive key derivations, we use SS-TWR without encryption but omitting the “report” message to prevent a neighbor to eavesdrops the “range” information. This choice affects the implementation of the ranging rendez-vous since the result will remain initiator-side. Consequently, the rendez-vous will be split into two ranging transactions: $A \rightarrow B$ and $B \rightarrow A$.

The rendez-vous time references are managed as follows:

1) The initiator of a TWR exchange schedules a request based on sent BLE advertisements timestamped in his local clock.
2) The responder schedules TWR RX windows based on scanned advertisement timestamped in his local clock.

The following safeguards are respectively added for TWR rendez-vous canceling per neighbor:

1) No BLE advertisement received in the last 60 seconds.
2) No successful TWR exchange in the last $TWR_{\text{expiry}}$ period.

Clearly, the initiator-responder rendez-vous window alignment will depend on the synchronization of corresponding BLE advertiser-scanner.

As per BLE specification, an advertised packet is broadcast in sequence on three channels (37, 38, 39). If the advertiser is notified when the third advertisement is sent (ADV END in Figure 4), the receiver is not aware of the channel on which he received the advertisement (3 possible events, see SCAN CH37/8/9 in Figure 4). It creates a small time hazard (up to 1.2ms in our PoC) that may result in the scanner being “early” to subsequent TWR rendez-vous. Therefore a large enough TWR receive (RX) window is needed, and we choose 2ns to account for being 1.2ms early, but also to encompass any OS delay in reporting ADV END events to the advertiser.

VII. EXPERIMENTAL EVALUATION

In this section we present the experimental results in terms of distance estimation accuracy (Section VII-A), power consumption (Section VII-B) and scalability (Section VII-C) we obtained using our PEPPER implementation on real hardware, in a controlled environment, and using our open platform.

![Fig. 4: BLE Scan and Advertisement Events](image)

TABLE I: Conducted experiments

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Test Mean</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>Consumption</td>
<td>fixed pair of nodes</td>
<td>Evaluate the effect of the number of ranging neighbors on a node’s current consumption</td>
</tr>
<tr>
<td>Accuracy</td>
<td>mobile pair of nodes</td>
<td>Evaluate the environment effect on the ranging accuracy under covid-relevant scenarios</td>
</tr>
<tr>
<td>Scalability</td>
<td>IoT-Lab, 14 nodes</td>
<td>Evaluate the UWB rendez-vous success probability under varying numbers of neighbors</td>
</tr>
</tbody>
</table>

![TABLE II: CT-relevant scenarios for proximity measurements](image)

Explored setups are listed in Table I and the main take-away is as follows:

1) **Accuracy**: looking at UWB and BLE performance through different obstacles, we can infer that UWB-based classification leads to fewer false negatives, in particular in presence of occlusion due to human activity (pocket, backpack, body). Although it comes with a higher false positive cost, it also means that fewer actual contacts are missed with UWB compared to BLE, which is a nice result from an epidemiological viewpoint.

2) **Power consumption**: the consumption of our PEPPER+DESIRE PoC increases linearly with the number of neighbors by $270 \mu A/\text{neighbor}$ approx. Although this behavior is close to expected, extra parameters tuning could further decrease energy footprint.

3) **Scalability**: With the default protocol configuration, we observed experimentally a rendez-vous success rate above 93% with up to 13 active neighbors.

A. Distance Evaluation Accuracy

In this section, we report measurements while carrying around the PEPPER+DESIRE token prototype in different scenarios depicted in Figure 5 and described in Table II. We deploy one static node and one mobile node (carried as a wearable shown in Figure 5). The experiments consist in 3 minute captures, over distances of $[0.5m, 1m, 1.5m, 2m, 2.5m, 3.5m]$, as we are mostly interested in accuracy when two tokens are near the critical distance for contact tracing, $d_c = 2m$. In order to increase the amount of BLE/UWB captures, devices were set in permanent scan and with advertisement rate of 5Hz.

![Fig. 5: Setups for CT-relevant scenarios](image)

The proximity data across all tests are shown in Figure 6. Focussing on the critical area below 2m, we can make the following comments:

9See the companion repository [24] for experimental artifacts.

10See the footnote on Presence Tracing (PT) in section VII-C regarding this number of 13 neighbors while considering scalability.
in case of a true contact, in presence of an occlusion due to human activity (i.e., pocket, backpack, body), the BLE classification is expected to fail due to fading and multi-path attenuation, causing RSSI values to drop significantly compared to the LOS reference situation. On the opposite, despite some dispersion, the average value of UWB ranging suffers less and remains close to the LOS situation, especially in the critical zone. From a CT perspective, UWB ranging is more resilient to false negatives in this case.

• in case of no contact because users have a physical separation (i.e., plexiglass, whiteboard), UWB still penetrates obstacles leading to a very accurate range estimation, whilst BLE RSSI gets highly attenuated compared to the LOS situation. From a CT perspective, BLE is more resilient to false positives in this case.

From an epidemiological viewpoint, reducing false negatives means that more actual contacts can be discovered, in situations where the use of BLE only could have wrongly concluded that the situation presents no risk. Having higher false positives is of lower importance and is the price to pay.

### B. Power Consumption

Using a Nordic PPK2 instrument, we can monitor the average current drawn by a DWM1011-DEV plugged as a Device Under Test (DUT) and configured with the PEPPER default parameters (see section VI). We check the effect of increasing the number of identically-configured neighbors along with the increase of the “TWREpiry Time” parameter. Figure 7 shows the consumption overhead w.r.t vanilla DESIRE:

- BLE (DESIRE): 1.68mA avg. draw or 5.587mW (3.3V);
- UWB overhead (due to PEPPER): 0.27mA avg. draw or 0.538mW (3.3V) per neighbor.

![Fig. 7: Power Consumption, linear regression ci = 95%](image)

For instance, under constant exposure to 16 neighbors, a 950mAh battery would last:

• 7 days with PEPPER: scanning at 25% rate, with 1 TWR exchange every 4 seconds (due to the 25% scan rate);
• 24 days with BLE-only DESIRE, same parameters.

The use of UWB negatively impacts autonomy, no surprise. However, setting $TWR_{Expire}$ to 30s, which is an acceptable tradeoff in front of 5 to 15 minutes long contacts, increases the battery lifetime up to 17 days.

### C. Scalability

Increasing the number of encounters increases the risk of missing a TWR rendez-vous. When a rendez-vous is scheduled, the worst case scenario yields an UWB transaction lasting $2 \times 3ms$: 2ms RX window, 760μs hold-off, 200μs air-time for the response, elapsed on each side. If we consider that when B opens a RX window for A, it cannot handle requests from other neighbors, then the expected average of non-overlapping rendez-vous can be modeled as a variant of the birthday problem:

$$E(V) = n \times \left(1 - \frac{1}{k}\right)^{(n-1)}, \quad k = \frac{1000ms}{2 \times 3ms}$$

In order to check how TWR exchanges success rate decays with the number of neighbors, we ran an experiment with all 14 DWM1001-DEV nodes on the Lille FIT-IoT-LAB. All the nodes were configured with default parameters (section VI) and we collected their Epoch data during 15 minutes. Results in Figure 8 show that when no $TWR_{Expire}$ policy is

\[11\] Scalability up to 13 neighbors may seem limited (e.g., in case of a crowded bar). However CT is complementary to “Presence Tracing” (PT) solutions (e.g. LUCA and CWA Event Registration in Germany, CrowNo-tifier in Swiss, CLEA in France). Specifically designed for public/private closed locations, these PT protocols rely on registering a person’s presence typically by scanning a dedicated QR code, and checking with a periodically updated black list of locations suspected to be cluster in certain time spans. The epidemiological assumptions between CT and PT are quite different, no study concluded on a superiority of one over the other to the best of our knowledge, both seem complementary (and have been used as such in the above countries), and PT has no scalability limit given its principles. With this in mind, our scalability experiments are reasonable in case of CT.
applied, the experimental values accurately fit the expected success rate in Equation 1. Indeed, a rendez-vous success rate above 93% for up to 13 neighbors is observed. This result is quite promising given that a high number of “active” neighbors is rather unlikely: DESIRE’s carousel advertising scheme requires devices to be in each other’s vicinity for at least 40s to trigger an encounter. On the other hand, adding a $T_{WR_{Expiry}}$ time for reducing TWR transactions does not significantly affect the rendez-vous success rate. Hence, this parameter introduces a good trade-off given the significant energy savings as shown in Section VII-B. However, when the number of neighbors is small, a discrepancy appears when $T_{WR_{Expiry}}$ is rather high. This effect can be explained by the fact that expiry-time also limits the number of exchanges per Epoch, and that the discovery process is asymmetric by nature. Indeed, missed TWR rendez-vous during discovery phase lowers the average success rate: A discovers B, but B has not yet discovered A.

![Fig. 8: Responder rendez-vous success rate](image)

**VIII. DISCUSSION & FUTURE WORK**

**Accurate Timestamps** — In Section VII we showed that UWB can be opportunistically activated based on BLE events. We also showed that accurate timestamps for advertisement and scan events are key in determining the required UWB RX window, which directly determines battery lifetime and scalability. From the advertiser viewpoint, OS and BLE stack delays in reporting the BLE advertisement-related events are key. As described in Section VI spacing in the three advertising channels will be the determining factor. But the BLE standard only specifies that advertisements on channels 37, 38, 39 are sent within at most a 10ms time frame. This could mean that a 20ms UWB RX window would be needed to cover the worst case, hindering battery lifetime and increasing rendez-vous collision probability (to be compared with our platform-related 1.2ms value measured, Section VII). A recent study [31] proposes a solution to detect the advertisement channel, but is not applicable to all devices, and also shows that many devices do not follow the standard advertisement channel pattern.

**Leveraging BLE 5.0 Features** — As DESIRE was designed to cover new and old devices alike, it does not rely on features present in recent Android releases (e.g., multi-advertising), or on BLE 5.0 features such as extended advertisement and periodic advertisement. These features could improve performance by offloading traffic to Secondary Advertisement channels with fixed advertisement intervals, allowing scanners and advertisers synchronization and therefore short BLE scan windows. This has a ripple effect on UWB synchronization, as advertisements happen on a single channel and can be more accurately timestamped, thus reducing the RX window size.

**Optimizing UWB for Contact Tracing** — So far UWB use cases mostly considered maximizing precision or range, transmitting at maximum power while respecting the international regulation (maximum average power spectral density (PSD) of 41.3dBm/MHz, and maximum peak power spectral density of 0dBm/50MHz). But CT scenarios only target a good accuracy within 3-4meters range. Therefore, transmit power, PRF and data-rates could and should be tuned according to this specific scenario: reducing time on air and power consumption.

Additionally, so far we have only discussed and analyzed the DW1001 HRP UWB chipset. This hardware is relatively old, and the DW3000 next generation consumes 40% less power. Even more, as recently shown [18], LRP UWB chipsets could be an energy and cost-effective alternative, reducing the power requirements tenfold.

**Decreasing False Positives** — In Section VII-A, we showed that UWB tends to exhibit more accurate classification in scenarios where BLE has so far been unable to perform, with a risk of a higher rate of false positives (e.g., with people standing on the opposite side of a wall). This could be alleviated by designing NLOS/LOS condition detection [32], or by using jointly BLE and UWB metrics in the classification and risk evaluation algorithms.

**Gauging the Attack Surface** — Privacy [33] has been identified as one of the main design requirements for CT applications, and has thus been their core concern [3, 4, 21]. In this paper, we have introduced the usage of UWB but have not revisited the risk analysis from DESIRE [20]. Accordingly, UWB higher precision could imply that through passive or active sniffing, attacker could precisely track devices, or perform DoS attacks. Some of these issues could be addressed with the use of silent periods [34], IEEE802154z [35] security, reduced TWR range through a reduced transmit power, TWR exchanges rate limitations, etc. The design of such countermeasures requires a specific analysis and is at this stage out of scope of the present work.

**IX. CONCLUSION**

This work introduces PEPPER, an innovative approach for digital Contact Tracing (CT) on disposable tokens. It leverages the privacy-preserving DESIRE contact tracing protocol, but also embedded and non personal tokens, RIOT as the underlying operating system for low-power IoT, and a joint use of Bluetooth Low Energy (BLE) and Ultra Wide Band (UWB) wireless protocols for reliable distance evaluations.

First, the full software is open source, which provides a very useful playground for further experiments in this domain.

Second, the unique flexibility of PEPPER in terms of deployment turns out to be a key advantage, since autonomous
tokens are not easily compatible with a decentralized risk analysis (e.g., GAEN requires each device to daily download megabytes of keys of COVID+ users, compute the 144 daily pseudonyms for each one, and compare them to recorded pseudonyms). By using concepts from DESIRE, PEPPER offers alternative deployment approaches and a unique privacy model that leverages per-encounter pseudonyms (rather than per-device with other CT protocols).

Third, this work also highlights benefits (e.g., UWB can decrease false negatives) and limits of jointly exploiting BLE and UWB radios for CT. We base these observations on experimental evaluations, using our prototype in different "everyday" environments.

This work opens the way for a mixed CT approach, where smartphone apps and dedicated disposable tokens could interoperate, opening new possibilities to end-users according to their preferences and possibilities (e.g., whether they own a compatible smartphone), while providing enhanced privacy guaranties.

References


