面向COVID-19大流行的基于LPWAN的新型健康码

**一、介绍**

新型冠状病毒COVID-19引起的肺炎2020年在全世界流行，造成了数以十万计的无辜生命死亡，致至少数千万人感染。世界卫生组织和各国政府采取了各种措施控制疫情的蔓延，但半年时间过去，仅有少数国家的疫情真正得到了控制，每天不断攀升的确诊和死亡人数触目惊心。人类在面临疫情时，显得格外渺小和束手无策，由于COVID-19远高于SARS和MERS的传染能力和高达14天的潜伏期，病毒在缺少防护的人群之间肆意传播，例如一个确诊患者在几小时的飞行途中就可以感染整个机舱的乘客，而这些被称为密切接触者的乘客又将病毒带下飞机，在工作单位和家庭中继续传播。更让人担忧的是无症状感染者的存在，这类人群感染了病毒，但却不表现丝毫的症状，仍以健康人的身份活动，直到一些偶然的机会才被发现，而在此期间已经接触了大量的人群。

与疫情初期各国防疫和医疗部门疲于救治重症确诊患者，而忽略了轻症患者以及大量的密切、间接接触者不同，随着医疗资源逐渐恢复正常，防疫经验逐步累积。目前世界上多国政府已经有能力在做好救治的同时，一旦发生新的局部疫情，就立即采取行动大规模对可能的接触者进行核酸检测，甚至对整个地区的人做检测。

传染病防治的一项重要任务就是流行病学调查，即追溯确诊患者在过去一段时间里（对于COVID-19一般是14天，一些国家甚至更久）的行踪，从而排查出他的多级密切接触者，是确定核酸检测目标对象的重要手段。流行病学调查主要依靠患者的自述，虽染目前很多地方也借助智能手机的传感器以及APP里诸如交易记录之类的信息进行调查，但是效果仍不理想，难免有漏网之鱼。6月，中国北京以农贸市场为中心的疫情之所以能被很快扑灭，得益于首位确诊患者的优秀的记忆力，他准确回想了自己过去14天的行踪，并提供了农贸市场这一关键信息，但专家似乎也很难考证患者的叙述是否完整和真实。另外，由于此次的疫情爆发出现在北京一个非常重要的市场，指向型极强，给后续排查接触者带来了一些便利。但事实上，这样的幸运事件并不总是发生，任何政府都不允许防疫工作建立在个人的运气之上，假设此次患者不能配合完成优秀的流行病学调查，或者患者停留的地方小而杂，恐怕就会给防疫工作带来很多不确定的挑战。因此，我们必须要拥有一套更好的追踪方法，不依靠个人的能力，与地点无关，准确安全，适合各类群体使用。

目前，各国政府和科技公司都在尝试提供科学的流行病学调查辅助方法。比较著名的是中国的健康码，各个省份甚至各个城市都推出了自己的健康码，这在中国被证明是一种有效的方式，因为在中国几乎所有的公共场所、小区都需要凭借健康码出入，因此一个参与社会生活的人必须申请，并自动纳入“联防联控”体系中。健康码大致分为三个登记，红色表示确诊患者或确诊患者的密切接触者和来自疫情高风险地区的人群，这类人在中国一般都住院治疗和由政府统一组织隔离，在社会生活中受到严格监管，不能出入任何公共场所，需要完成至少14天的隔离或治疗。黄色表示从疾病中康复的人员、确诊的病患的一般接触者或来自中风险地区的人员，这类人群根据属地的政策一般需进行7天的隔离观察，在社会生活中也受到比较严格的限制。绿色表示完全健康的人群，也是绝大多数人群的健康码颜色，指的是完全没有疫情接触史的人群或者得到了医疗机构认可从红黄码中隔离结束的人群，绿码持有者可以无障碍的在属地活动，不应受到歧视或阻拦。健康码的使用流程，以作者所在的辽宁省为例，一般是，首先申请者要拥有一台智能手机，下载所在地政府所规定的APP或者小程序，按照里面的要求填写个人信息以及疫情有关的信息，包括是否去过疫情高风险区，是否接触过确诊患者，以及自身是否有疫情相关症状等，之后系统根据这些信息生成对应颜色的码。使用的办法一般包括两种，一是向工作人员出示健康码以进入一些公共场所，更为有效的办法是用手机扫描贴在场所处的二维码，从而不仅可以验证现有健康码的颜色也可以将行踪进行记录。这种方法在中国大面积推广且取得了一定的效果，在防疫过程中起到了重要作用。然而，它存在几个非常重要的缺陷。第一、许多老年人、儿童没有智能手机或者即使有也在操作上有很大的障碍，据报道，截至2020年3月，中国网民规模为9.04亿，互联网普及率达64.5%。也就是说，中国大约有5亿人是不上网的，除此之外，在2019年中国60周岁及以上人口有25388万人，这其中，接触过网络的只有6056万人，而能熟练使用智能手机的占比更是少之又少，这还没有考虑到尚不能独立使用电子设备的儿童，即使填报流程可以由他人代劳，但是在使用时动作缓慢，堵塞在公共场合门口，极大影响了后面的人流，这时往往会采用手动登记或者录音的办法，从而增大了社交风险，为病毒传播提供了额外的机会。第二、由于工作人员不足，或是忙于测温工作，导致对健康码的检查存在漏洞，主要包括，（1）是否有码：部分人没有健康码却跟随人流混入场所，也存在因为手机没有电量导致确实不能使用的情况（2）码的时效性：一些人为了图方便出示截图（3）出示码而不扫码：这是在实际操作中最常出现的漏洞，这只能证实当前的健康状态，而无法在疫情发生后进行追溯。第三、由于网络环境导致APP无法调用，从而滞留大量人群引发聚集风险。第四、由于各地政府的政策不一致，导致每到一个地方就要申请新的健康码，甚至同一人到不同的省市申请到的健康码结果不同的情况，而在健康码互认的问题上更是矛盾多多，据报道多次出现了武汉市民持有绿码却在外地寸步难行的尴尬情况。另外，在用户使用的友好程度上，现行的健康码也并不令人满意，首先就是一次出行需要反复的打开手机，而由于佩戴口罩又不能快速解锁手机，再调用软件扫码和展示码，这一过程每天可能要占据10余分钟的时间。其次，APP的安全性令人担忧，虽是由政府掌控个人关键信息，但是难以保证其他恶意软件通过手机上的漏洞窃取到健康码APP的内容。如果说以上提到的还都是一些小的缺点和不足，那么由于智能手机性能、APP权限、网络故障和人为因素导致的关键信息遗漏问题就可能大大影响流行病学调查的准确性和完整程度，进而导致十分依赖于此的防疫工作功败垂成。

而从三月起就成为疫情新的震中的美国，截至7月中旬确诊病例已经接近400万，并且丝毫没有放缓的趋势，COVID-19已经在美国肆虐开来，而由于文化习惯，美国民众似乎并不喜欢佩戴口罩，这给疫情传播提供了极佳的条件，也为确诊患者的密切基础者追踪带来了极大的困难。因此，五月份两大科技巨头苹果和谷歌联合退出了一种基于蓝牙的密切接触者追踪系统，工作原理仍基于智能手机和APP，且要求设备的蓝牙保持开启，系统运行时，手机就会周期性地弹出一段小的、唯一的、匿名的代码，这段代码是从手机的唯一ID中提取出来的。其他手机会接收并记住这段代码，建立一个记录他们接收到的代码和接收到它们的时间的日志。相反，设备也会接收其他设备发来的代码。为了保护隐私，这段代码还会每隔一段时间更新一次。当使用该系统的人确诊患病后，他可以选择将自己的确诊信息上传到服务器上，其他用户定期从服务器上下载确诊信息，并与本地所存储的他人代码比对，如果有匹配项，则意味着自己与确诊患者曾密切接触过，从而决定自己的下一步医治计划。此类系统在隐私保护的原理上似乎没有漏洞，可能引起的隐私泄露问题尚在讨论，但单就其防疫效果来讲是很差的。首先，要求这些甚至反对口罩的美国民众下载此APP并需要打开蓝牙功能就是很困难的，出于文化和历史的原因，愿意花时间和精力去做防疫工作的民众并不占绝对多数，如果不是参与社交的每一个人都开启了这项服务，那整个过程的效果就大打折扣，极端来讲，在一个社交场合恰好有一个人没有开启服务，而恰恰这个人是确诊的患者，则其他所有人的参与都是徒劳。第二、依靠患者自主上传的方式并不可靠，确诊者极有可能因为主观或是客观原因没有上传确诊信息，那整个环节就失效了。第三、即使前序流程顺利，当一名用户获知自己成为了密切接触者时，他极有可能因为是一名无症状感染者或仍存在潜伏期没有发病而拒绝前往医疗点检测和治疗，并继续参与日常社交活动传播病毒。可怕的是，由于此系统的保密机制，只要用户不前往医院，外人甚至是政府医疗机构无法获知用户的身份和接触情况，也不能及时为其提供最佳救治方案。如果面对的是比COVID-19更致命的病毒，其中一环的缺失就将引起巨大的后果。如果再综合考虑其本身的工作机制和无线通信技术，该方法更是存在耗电较大、不同机型间的融合以及蓝牙传输距离限制导致无法获取稍远一点的接触者信息的问题，据报道，病毒不仅可以通过已知的飞沫方式传播，也可能通过气溶胶传播到稍远一些的地方，而这个距离很有可能已经超过了蓝牙的传输范围。此外，该方法仅限第一级密切接触者的追踪，而无法掌握次一级密切基础者的密切接触者，需知在各国报道的病例中，存在不少跨层的传染途径。

其他的方法还包括借助运营商网络的，比如中国国务院、信通院与三大运营商合作的防疫大数据行程码，可以根据手机使用连接基站的情况判断持有者到访过哪座城市，甚至哪个区。此码的作用更多体现在远程出行中，并不针对特定的确诊患者的密切接触，而是从地区的风险评级上判断一个人可能的健康程度。与健康码相比，此系统似乎不需要用户的主动参与，但是在实际使用中其灵敏度不高，例如短暂停留一个城市时，并没有该城市的记录，而需要至少拨通一次电话才可以。相反，也出现了仅仅在火车上路过却被记录在案的情况。另外一个风险是，一些投机者可以利用双卡的优势，完美的避开轨迹被记录，从而给追溯带来困难。还有一些地方的政府设计的被动的基于位置的追踪系统，这种方法也作为一部分包含在前文所述的中国健康码中，APP利用手机的GPS、Wi-Fi等功能，对手机进行定位，从而获取持有者的轨迹，但是无线定位技术本身存在的误差和场景局限性是制约此方法的障碍，而且只要其依靠智能手机工作，就仍然逃不掉被其他恶意软件窃取信息的威胁。一旦手机电量耗尽、用户私自关闭了某些功能，都将导致数据收集和疫情追踪的失败。

我们无法仅以此判断哪种追踪措施更好，因为在不同文化背景下生活的人群会对同样的事情有截然不同的看法，因此也会产生差异化的应对方式，就像对待是否应该佩戴口罩这件事，似乎在一些国家已经争论了半年多的时间仍然没有结论。本文抛开文化观念问题，完全从应对COVID-19甚至将来有可能发生的更为严重的传染病疫情角度出发，试图提出一种新的密切接触者追踪方法，我们认为一个好的解决方案应该包含以下特点：

1. 简单易用：我们希望新的办法可以降低对使用者的能力要求，无论是老年人还是儿童，要求尽可能少的操作；
2. 低耗耐用：如果必须依靠智能手机，应当尽可能降低因电量耗尽带来的尴尬，或许应该有更高效且低成本的替代品；
3. 不得不用：新的方法要所有人的参与，无论是出于对自身的防护还是对家庭社会的责任，你应该像佩戴口罩一样使用它，政府和医疗机构也需要始终监督，从而第一时间对有风险得人群进行检测和隔离；
4. 可靠好用：确诊患者接触信息是流行病学调查的重要依据，必须要保证数据的可靠性，任何人不能借助漏洞规避追踪，不能擅自脱离系统或是对已生成数据进行篡改；
5. 通用：一套方法适用整个地区，不需要换一个地方换一个方案。

因此，我们设计了一个基于LPWAN远距离通信技术和区块链技术的实体健康码设备，可以满足以上的所有要求。首先，仅需要在领取时进行一次登记即可，将个人的ID、联系方式信息录入，今后的使用中不需要任何操作。其次，依赖于LPWAN的超低功耗，一个设备可以工作数年以上，不仅可以供应对COVID-19使用，即使不幸遇到新的疫情，也可以拿出来继续使用，而且单个设备的成本非常低廉，体积足够小且可以根据需求制作成钥匙、卡甚至挂饰，并不给用户带来忘记携带难以携带的困扰，方便大规模推广。我们吸纳中国健康码的设计，为设备设置了同样三种颜色的指示灯，再也不需要进行拿出手机-解锁-扫码-展示的复杂环节，只需要拿出设备，工作人员就可以看见你的颜色并放行，再也不需要担心老人或者儿童不会使用，并确保了每个人的参与。我们的设计是密闭的，仅记录了必要的联络信息，且由于他的工作不需要依靠任何其他智能设备，避免了被窃取的可能性。另外，由于没有任何可修改操作的可能性，只要携带设备参与了社交活动，你与确诊患者的接触记录就会被储存。我们还采用了区块链技术防止数据被篡改，向任何想徇私枉法影响防疫工作的人说不。最后，这套方案有着很好的通用性，不需要申请多个健康码，用户就可以畅通的出行到部署此系统的地区。

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 方法名称 | 国家/机构 | 主要技术 | 能耗 | 覆盖 | 精度 | 主要功能 | 技术安全性 | 隐私保护 | 易用性 | 可靠性 |
| 中国健康码 | 中国 | 二维码，GPS | 中等 | 非常广 | 较高 | 到访地点 | 比较高 | 较低 | 中等 | 需要较强的人工监管力度 |
| 行程码 | 中国 | 蜂窝网络 | 较高 | 非常广 | 低 | 到访城市 | 比较高 | 较低 | 低 | 有一定漏报率 |
| TraceTogether | 新加坡 | 蓝牙 | 高 | 低 | 高 | 密切接触者 | 低 | 中等 | 中等 | 依靠用户自觉性 |
| Google/Apple Contact Tracing | 美国 | 蓝牙 | 高 | 低 | 高 | 密切接触者 | 低 | 高 | 中等 | 依靠用户自觉 |
| BeepTrace | University of Glasgow | GPS，蓝牙、蜂窝网络、WiFi、区块链 | 高 | 高 | 高 |  |  | 高 | 中等 |  |

**二、背景知识**

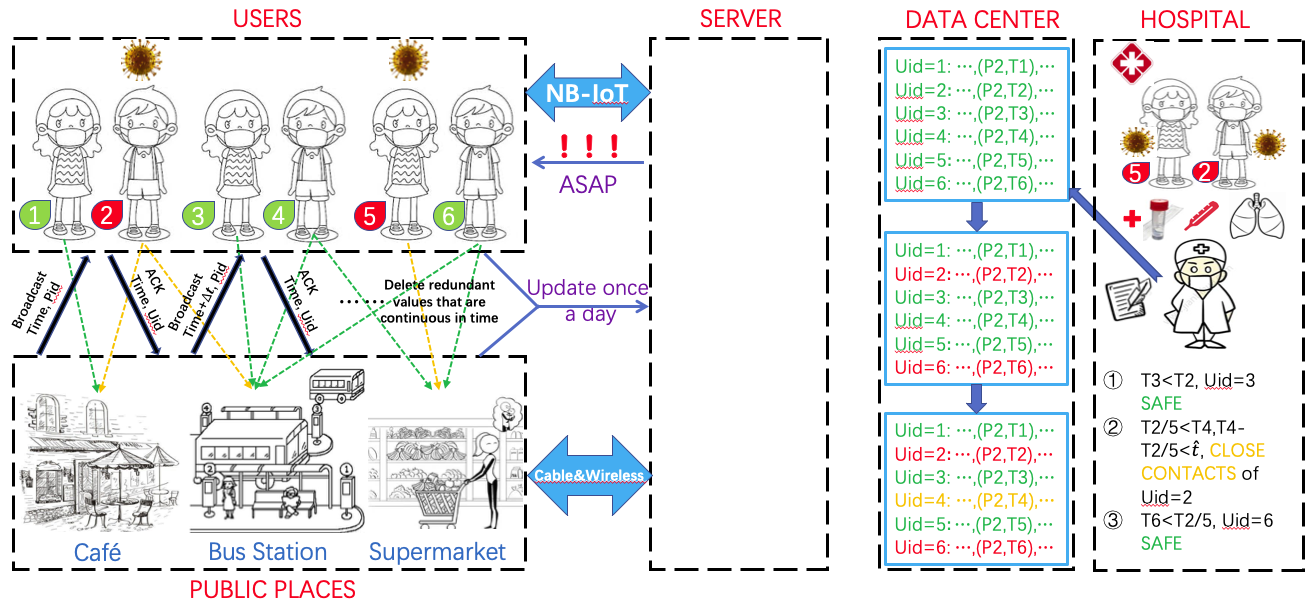
物联网通信技术繁多，从传输距离上可划分成两类：第一类是短距离通信技术，例如ZigBee、Wi-Fi、Bluetooth等，典型的应用场合如智能家居；第二类是低功耗广域网(Low Power Wide Area Network, LPWAN)，典型的应用为智能抄表系统。LPWAN技术又可根据工作频段分为两类：一类工作在非授权频段，如LoRa、SigFox等，此类技术无统一标准，也没有大型运营商提供支持，功能、规模完全自定义实现；第二类工作于授权频段下，3GPP支持的2/3/4G蜂窝通信技术，如全球移动通信系统(Global System for Mobile Communication , GSM)、长期演进(Long Term Evolution, LTE)和基于蜂窝的窄带物联网(Narrow Band Internet of Things, NB-IoT)等。

NB-IoT只消耗大约180KHz的带宽，可采取带内、保护带或独立载波等三种部署方式，可直接部署于GSM网络、UMTS网络或LTE网络，重用现有的网络资源。在所有主要移动设备、芯片组和模块制造商的支持下，NB-IoT可以与2G，3G和4G移动网络共存。它还受益于移动网络的所有安全和隐私功能，例如对用户身份机密性，实体身份验证，机密性，数据完整性和移动设备标识的支持。因此基于NB-IoT开发的疫情追踪系统相比使用其他技术可以降低或免去部署成本、实现平滑升级。

NB-IoT支持待机时间长、对网络连接要求较高设备的高效连接。NB-IoT设备电池寿命可以提高至至少10年，同时新的物理层信号和信道旨在满足扩展覆盖范围以及超低设备复杂性的苛刻要求，从而提供非常全面的乡村和室内蜂窝数据连接覆盖。NB-IoT模块的初始成本远低于其他类型的无线通信技术，在LPWAN领域内也有一定竞争力，单个模块的成本在2-5美元左右，并且该基础技术因依赖于蜂窝网络比LoRa，SigFox更为扎实，并且随着需求的增加，其平均成本预计会迅速下降。

在基站数量相同的情况下，NB-IoT可以比现有无线技术提供高达50~100倍的接入数。一个扇区能够支持10万个连接，支持低延时敏感度、超低的设备成本、低设备功耗和优化的网络结构。

本系统的工作流程如下



首先，每个人需要领取一个智能追踪设备，一个普通的设备包括（1）ZigBee收发端（2）NB-IoT收发端（3）共享的存储空间（4）可以显示三种颜色的LED指示灯（5）电池。该设备可以根据需求定制为可穿戴、卡片或是其他形状。每个设备有其特定的号码，即NB-IoT接入运营商网络的号码，例如在中国，三大运营商会为你提供一个唯一的11位号码，就像手机号码一样。申领设备可以通过社区、学校、社会组织或是运营商等渠道，因为不需要面对面的设置，甚至可以通过邮寄的方式无接触完成。

其次，在拿到设备后，用户需要激活并注册个人信息，在世界上不同的地区，区分市民的方式不尽相同，为了保持一贯性，本文后续的描述均以作者所在的中国大陆地区为例。如果本文后续介绍的追踪系统确认用户为具有较高风险确诊患者的接触者，政府部门可以通过预留的信息联系到该用户，如果被断定为具有一般风险的接触者，用户也可以从设备本身的指示灯上获知，从而可以选择自行前往医疗机构进行检测或隔离。在另一端，即场所端，负责人仅需要设置好场所的信息，配以统一管理的场所ID即可，根据公共场所的面积大小不同，可能需要设置多个设备并共享同一ID，由于场所的设备处于长期工作状态，因此需要配备电源，并不难实现。需要特别注意的是，本文特别考虑了现有工作中忽视的小型公共场所的追踪问题，例如一些临街小型饭店、便利店或私人商铺等，在7月中国大连的疫情中，首个感染者正是频繁活动于一个位于社区中的小型棋牌室，导致一些同处的市民被感染，这类场所人流量大且人员成分复杂，很难追踪。受限于现有无线技术的精度，在一定功耗限制的情况下，很难实现如此精确的定位，扫描场所二维码的方法虽然有效，但受限于工作人员数量，很难真正监管到位。因此，本文创造性的提出将追踪设备两用，对于这类场所的经营者，可以将个人的追踪设备作为小型场所的公用设备，即个人使用时与一般用户一样作为信号的接收方，而在从事个人经营活动时，切换使用模式将设备作为信号的发送方，从而在不需要大规模增加设备的前提下，将小型场所覆盖入追踪系统。据报道，超过半数的聚集性疫情发生在饭店等中小型公共场所，由于信息的缺失导致追踪不力，造成了非常大的损失，此方案的提出将很好的弥补这一漏洞。

在实际使用过程中，整个系统与用户无关，不需要用户进行任何操作，这将摆脱现有的系统对智能手机的强烈依赖，从而使得感染风险更大却被使用门槛阻拦的老年和儿童群体纳入到健康安全监控体系。其内部运行方式是：对于一个公共场所的设备，始终以周期性发送ZigBee信号，每个数据包仅包括场所ID和时间(Pid, time)，用户的设备收到后反馈ACK，并加入用户ID(ack, Uid)。通过对大量新闻报道中对确诊患者感染过程的描述并综合考虑设备的能量消耗和存储能力，我们将发送间隔设定为Δ=15分钟，则对于一个用户，每天最多收到96个数据包。而为了可以进一步降低用户设备的负载，我们设计了对数据包的舍弃机制，如果一个设备连续收到包含同一Pid的数据包，则其只保留首个包P0和最后一个包Pn，并计算停留在该处的时间段为首个包P0\_time-Δ至Pn\_time+Δ，因为如果不巧一个用户恰好是在某个包刚刚发送结束后进入而又在某个包发送前离开，就可能最多少计算了2Δ的时间。当仅收到一个包时，也将计算至少2Δ的接触期。考虑到ZigBee本身传输的丢包，尽管我们尽可能密集的发送数据包，仍无法保证百分之百的接收率，因此，如果在收到一个来自Pid=m的包后，相隔大于等于2Δ而小于等于4Δ的时间内（即30至60分钟）再次收到了Pid=m的数据包，并且在这段时间内没有收到来自Pid≠m的信号，我们认为该用户始终处于Pid=m的场所，并按照上文所述取第一个和最后一个包计算停留时间。基于这样的办法，可以将因无线信号本身性能带来的对整个系统的影响尽可能的降低。用户设备将记录这些值，并存储到本地。

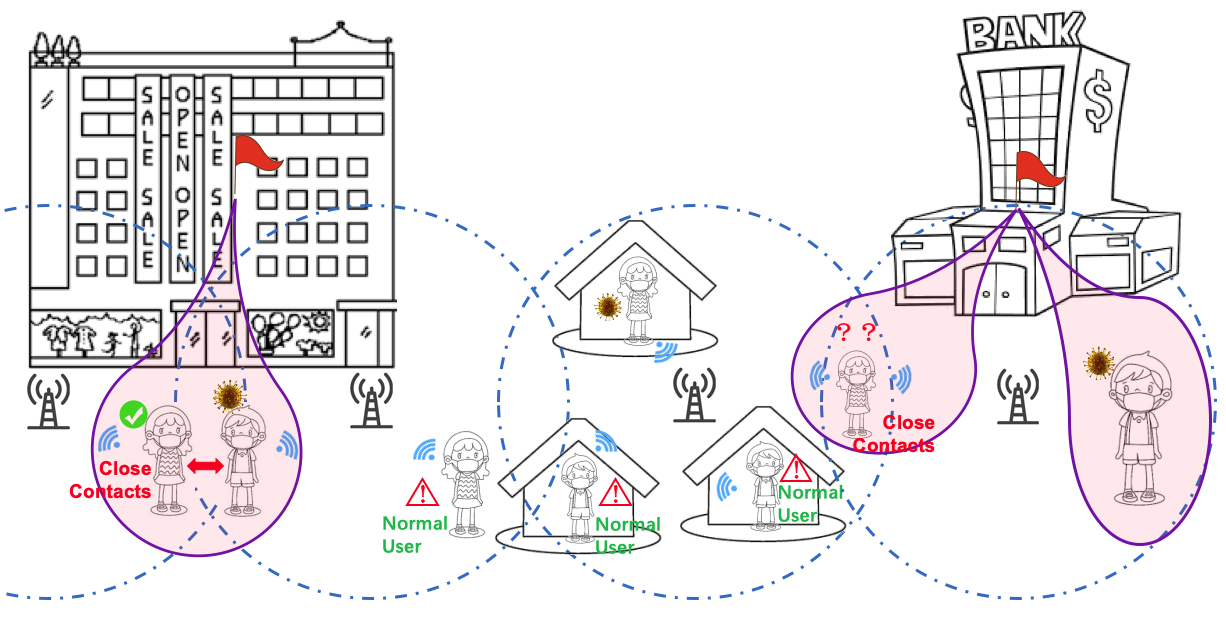
如何确保用户收到了信号？

由于疫情的追溯均以天为单位，在COVID-19中为14天，因此在数据上传的过程中本系统将以天为单位上报给服务器，在本地，以时间为参数封装一个或多个数据包。为了缓解同时上传数据带来的负载问题，次日，所有用户将通过NB-IoT端分批的向最近的基站上传前一日的数据。因为疫情的追溯并不是一个实时性要求特别高的应用，最大24小时的延迟是可以的接受的。当一名用户被确诊需要进行追溯而数据

没有完全上传时，管理员也可以通过发送下行的命令要求设备立刻补全数据。

值得注意的是，我们在本系统的不同阶段采用了两种不同的传输技术，这显然会比使用单一技术更复杂。然而，NB-IoT技术复用了运营商的基站，而在5G的进程中，在很多一线城市5G基站的密度也仅能达到200米一个，且大部分在室外，这样的部署密度显然难以区别不同的室内环境，而疫情追溯一般不考虑室外接触，如果在用户和场所交换信息时采用NB-IoT技术，虽然速度和连接性可以保证，但是这样不仅在关键区域有较大误差，却又在不关注的区域浪费了覆盖范围，如图所示是几种误判的情况：

1. 一个大型建筑物由两个相邻的基站覆盖，一对密切接触者恰好处在两个不同的基站范围中，因此没有被正确判定为重点人群；
2. 一位用户出现在确诊患者所连接的基站范围内，然而其身处室外并没有密切接触，被误判断为重点人群；
3. 多个小型建筑物在同一基站范围内，其中一个小型建筑出现了确诊患者，而在其他建筑物内完全没有接触的两位用户被误诊为重点人群；
4. 一位用户与确诊用户在室内接触，而恰好处在两个基站交界区域，因为信号的不稳定导致连接经常切换，无法准确判断是否为重点人群。



而ZigBee的通信范围在十米到几百米之间，可以根据需要调节发送端的能量值来控制传输的范围，因此也被广泛用于智能家庭的部署之中。采用ZigBee技术作为用户与访问地点之间交互的技术，大型建筑物可以通过多个收发端的组合实现全覆盖，而小型场所也可以仅使用普通的用户终端作为发射器，并调节功率达到覆盖。其不能作为用户与数据库之间数据传输的桥梁的原因显而易见，一是需要为其额外部署设备，成本比较高，二是其在城市内明显的衰减导致其传输范围进一步减小，不满足该场景的需求。

我们也在前期调研工作中考虑了很多其他的技术，室内环境Wi-Fi可以依赖智能电话避免新增设备，但是其能耗较高，并存在安全和使用门槛问题。而在室外LoRa技术似乎是一种可以考虑的方案，其传输范围足够长，低能耗也满足长时间的使用，但是缺点在于LoRa工作在未授权的频段，为单一的应用大面积部署网关性价比不高，在没有运营商维护的情况下通信质量也难以保证。采用蜂窝网络的传输仍无法避免对智能手机的依赖，且其能耗过大，不适合用于可长期使用的可穿戴设备。