Kevin Curran Professor of Cyber Security
kj.curran@ulster.ac.uk
Ulster University, Derry

The security implications of hosting public-facing health records in online services

On 2 July 2018, the UK Government announced a new National Health Service ('NHS') app that will put patients in England in direct touch with their general practitioners ('GPs')[1]. It is designed to enable users to book appointments, make repeat prescriptions and view their personal medical files that are held by surgeries. Testing of the app will begin in the autumn, and it should be downloadable in England by December 2018. Kevin Curran, Professor of Cyber Security at Ulster University, discusses the security risks that may surface alongside the benefits of such technology being rolled out to the public.

This app is interesting in that simple tasks, such as booking appointments and making repeat prescriptions, can be done electronically quite easily. Some patients have been able to book appointments and order repeat prescriptions through their local health centres' websites for several years. However, what is new here is access to medical records via an app. There is always a worry about sensitive information that is viewable within such an app potentially being viewed by others who gain physical access to a user's phone - hence the need for properly securing it with a strong personal identification number ('PIN') or biometric authentication. However, the main worry is the NHS's back-end system, which provides the application programming interface ('API') (or data) to the outside world, and which is what cyber criminals would hit. Any systems that provide external-facing data must be 'bullet proof' in their authentication mechanisms and have in place a myriad of other protections to prevent the huge list of critical security risks to web applications.

The Open Web Application Security Project ('OWASP') is an open community dedicated to enabling organisations to conceive, develop, acquire, operate and maintain applications that can be trusted. It has a famous web application security risk top 10, which lists the top web vulnerabilities[2]. In fact, on the same day that news of the NHS app was being announced, it was revealed that a coding error had led to 150,000 patients in England being involved in a data breach of NHS records[3]. Developing secure, robust web applications in the cloud is hard. Even those developers who understand secure coding also need to understand how to encrypt databases, prevent structured query language ('SQL') injection attacks, know about third-party library vulnerabilities, ensure that passwords are hashed, implement multi-factor authentication, prevent denial of service attacks, ensure that no resources are enumerable in the public API, do client-side input validation, know how to configure cloud services, and isolation of processes, use HTTP strict transfer security ('HSTS'), use intrusion detection systems ('IDSs'), patch underlying virtual machines, restrict ports and ensure minimal access privileges - and that is only for starters. A hacker needs to find only one flaw that allows them access, whereas the NHS systems administrators must ensure that every known vulnerability is patched.

In highlighting this point, I am not trying to stop technological progress: more than anyone, I can see the benefits of accessing information remotely. However, there is data and there is 'data': very little data is more personal than medical records. I also understand where moves towards allowing access to such records via an app have come from. I am sure that, over the years, there have been many meetings with senior management in the NHS, at which lobbying was done to move everything onto public-facing websites. I am sure that the argument was made that the information was already stored on computers. However, there is a big difference between having computerised records within the NHS information technology ('IT') infrastructure and having them reside on a public-facing server. Having such records within an NHS infrastructure limits the range and type of access that can be made. There are breaches, but the level of difficulty for remote hackers is much higher. Many such systems are isolated, with records from certain healthcare trusts not being linked to those of others. There are also legacy systems and non-standard technological platforms.

The problem with moving all records to a cloud-based API (we presume that it will be on the cloud somewhere) is that we will end up with a central focal point at which hackers will be able to aim all their arsenal and chip away until the door comes down and all data is available. Unprotected databases are very easy to find. Criminals use network visibility tools such as the freely available search engine Shodan, which indexes internet-connected devices. If we were to search Shodan right now for exposed databases, using a database term such as 'MongoDB,' we would find over 65,000 exposed databases. Of course, they may not all be vulnerable to attack, but simply being visible increases their risk of being breached. Health records are quite lucrative to criminals, due to their being rich in personal data, such as our social security numbers, medical histories, addresses, next of kin, dates of birth and, in many countries, insurance information and credit card details.

There are models that the NHS can adopt in order to limit future data breaches. A simple one is to make

1. Cellan-Jones, R., 'NHS app: Will it cut down on wasted appointments?' BBC News, 2 July 2017: https://www.bbc.co.uk/news/technology-44676493
2. OWASP Top Ten Project, 3 July 2018: https://www.owasp.org/index.php/Category:OWASP_Top_Ten_Project
3. 'NHS data breach affects 150,000 patients in England', BBC News, 2 July 2017: https://www.bbc.co.uk/news/technology-44682369
4. Felton, E. and Kroll, J., 'Heartbleed shows government must lead on Internet Security', Scientific American, 1 July 2014: https://www.scientificamerican.com/article/heartbleed-shows-government-must-lead-on-internet-security1/
5. McBrearty, S., Farrelly, W. and Curran, K., 'The Performance Cost of Preserving Data/Query Privacy Using Searchable Symmetric Encryption', Security and Communication Networks, Vol. 9, No. 18, 5311–32, DOI: 10.1002/sec.1699, Wiley.

continued

the system 'opt in,' so that those who decide that the positives outweigh the negatives can choose to have their medical information moved to a public-facing service so that they can access it. However, those who do not opt in or download the app and use it should, by default, have their records hosted in a non-public-facing cloud service. That way, if a data breach does occur, those who have never used (or wanted to use) the app will not have had their details released. This is not rocket science, and it would be trivial to implement. Basically, we would be looking to limit the data set. There are some who believe that governments should step in and lead on internet security, given the importance of the internet in modern life, but that is not an obvious solution[4].

Another option available to the NHS would be to use a form of fully homomorphic encryption ('FHE') which supports computations over data in encrypted form, including searchable symmetric encryption ('SSE') as it was originally envisioned; nonetheless, efficient FHE remains some way off[5]. In a cloud environment, cryptography is typically utilised for two purposes: security while data is at rest; and security while data is in transit. Unfortunately, the cloud cannot guarantee the security of data during processing, as the current limitations of cryptography prevent it from being processed in encrypted form. Given that data is processed in unencrypted form, it is quite common for attackers to target data in use, rather than targeting data that is encrypted during storage and transit. That is where modern techniques such as FHE, oblivious RAM ('ORAM') and searchable encryption could be considered. Used in isolation, ORAM does not support searchable encryption. Essentially, ORAM is a client-server communication protocol that is designed to obfuscate memory access patterns on the server side of a given transaction. In the context of searchable encryption, ORAM is typically combined with SSE and public-key encryption with keyword search ('PEKS') searchable encryption schemes to improve their security. SSE and PEKS searchable encryption schemes leak information to the server a few ways. By combining such schemes with ORAM, such information leakage can be eradicated. Nonetheless, the search efficiency of schemes utilising ORAM is severely hindered due to the amount of work involved in obfuscating memory access patterns using it. In relation to search efficiency, both SSE and PEKS achieve optimal search time when used in conjunction with an inverted index; that is, search time is linear in the number of documents matching the search string. However, on security, SSE is vastly superior to PEKS. Given that PEKS is a form of public-key encryption, an adversary can easily mount an attack on such a searchable encryption scheme, given the associated public key and a dictionary of chosen terms. In the case of SSE, all associated keys are kept private. SSE represents one of the few forms of searchable encryption that are achievable using established standardised encryption algorithms. Alternative forms of searchable encryption require the use of non-standardised, special-purpose encryption algorithms. SSE is considered one of the least secure forms of searchable encryption, primarily due to information leakage. There exist solutions to eradicate and obfuscate all forms of information leakage in SSE; however, existing solutions have a significant effect on its search efficiency. Evidently, the challenge for organisations such as the NHS is to improve the security of SSE while maintaining its superior search efficiency.

Of course, the technical teams involved in rolling out the NHS system will strive to deliver a secure service. They may use best practices, such as penetration testing (which is common for probing systems) but many unintentional, yet significant, security problems cannot be found through penetration testing alone, and therefore source code auditing is the technique of choice for technical testing. Auditing code manually can be particularly effective for discovering issues such as access control problems, Easter eggs, time bombs, cryptographic weaknesses, backdoors, Trojans, logic bombs and other malicious code. Reputable organisations will execute internal code audits, but in order to gain wider acceptance they also need to invoke external code audits that will give (or not) external validation of a product's ability to meet expected requirements (for example, integrity, confidentiality and availability).

The fact that the new service utilises an app means that, for the most part, users of mobiles should follow the same safe computing principles as they would on traditional desktops. Many mobile service providers have security policies in place such as secret questions or personal PINs, along with multi-factor authentication. Users should always set passcodes, keep them locked when not in use and use biometric features if they are available. They should not store personal details such as passwords or PINs in texts or emails on the device. They should also be aware of the problem of rogue networks and the possibility of 'man in the middle' attacks on public Wifi networks. They can install mobile anti-virus clients, paid and free options for which are on the market, but, unfortunately, some of them can be 'heavy weight,' taking a toll on overall device performance and battery life. Again, security often comes at the expense of convenience.

To conclude, to date, the relatively short modern history of IT has already shown us that organisations are not good at securing access in web-facing portals, so the decision to place such sensitive information online at this time is certainly interesting. The issue with accessing health records via a mobile app is not so much about security but rather that it will offer a public-facing service that will provide health records. The world's information is moving online, and people expect to be able to access remote services globally 24/7, but we should be aware that there is personal data and there is also person data. Health records are in the most sensitive category of personal data, so it will be expected that military-grade encryption and protection and the strongest of user authentication mechanisms will be in place. Does the NHS believe that it has those? I certainly hope so.