

## *Insecure by design: protocols for encrypted phone calls*

*Steven J. Murdoch*

*March 2016*

The MIKEY-SAKKE protocol is being promoted by the UK government as a better way to secure phone calls. The reality is that MIKEY-SAKKE is designed to offer minimal security while allowing undetectable mass surveillance, through the introduction a backdoor based around mandatory key-escrow. This weakness has implications which go further than just the security of phone calls.

THE CURRENT STATE OF SECURITY FOR PHONE CALLS LEAVES A LOT TO BE DESIRED. Land-line calls are almost entirely unencrypted, and cellphone calls are also unencrypted except for the radio link between the handset and the phone network. While the latest cryptography standards for cellphones (3G and 4G) are reasonably strong it is possible to force a phone to fall back to older standards with easy-to-break cryptography, if any. The vast majority of phones will not reveal to their user whether such an attack is under way.

The only reason that eavesdropping on land-line calls is not commonplace is that getting access to the closed phone networks is not as easy compared to the more open Internet, and cellphone cryptography designers relied on the equipment necessary to intercept the radio link being only affordable by well-funded government intelligence agencies, and not by criminals or for corporate espionage. That might have been true in the past but it certainly no longer the case with the necessary equipment now available for \$1,500<sup>1</sup>. Governments, companies and individuals are increasingly looking for better security.

A second driver for better phone call encryption is the convergence of Internet and phone networks. The LTE (Long-Term Evolution) 4G cellphone standard – under development by the 3rd Generation Partnership Project (3GPP) – carries voice calls over IP packets, and desktop phones in companies are increasingly carrying voice over IP (VoIP) too. Because voice calls may travel over the Internet, whatever security was offered by the closed phone networks is gone and so other security mechanisms are needed.

Like Internet data encryption, voice encryption can broadly be categorised as either link encryption, where each intermediary may encrypt data before passing it onto the next, or end-to-end encryption, where communications are encrypted such that only the legitimate end-points can have access to the unencrypted communication. End-to-end encryption is preferable for security because it avoids intermediaries being able to eavesdrop on communications and gives the end-points assurance that communications will indeed be encrypted all the way to their other communication partner.

<sup>1</sup> Kristin Paget, "Practical Cellphone Spying", blog post, 2010, <http://www.tombom.co.uk/blog/?p=262>

Current cellphone encryption standards are link encryption: the phone encrypts calls between it and the phone network using cryptographic keys stored on the Subscriber Identity Module (SIM). Within the phone network, encryption may also be present but the network provider still has access to unencrypted data, so even ignoring the vulnerability to fall-back attacks on the radio link, the network providers and their suppliers are weak points that are tempting for attackers to compromise. Recent examples of such attacks include the compromise of the phone networks of Vodafone in Greece (2004)<sup>2</sup> and Belgacom in Belgium (2012)<sup>3</sup>, and the SIM card supplier Gemalto in France (2010)<sup>4</sup>. The identity of the Vodafone Greece hacker remains unknown (though the NSA is suspected<sup>5</sup>) but the attacks against Belgacom and Gemalto were carried out by the UK signals intelligence agency – GCHQ – and only publicly revealed from the Snowden leaks, so it is quite possible there are others attacks which remain hidden.

Email is typically only secured by link encryption, if at all, with HTTPS encrypting access to most webmail and Transport Layer Security (TLS) sometimes encrypting other communication protocols that carry email (SMTP, IMAP and POP). Again, the fact that intermediaries have access to plaintext creates a vulnerability, as demonstrated by the 2009 hack of Google’s Gmail<sup>6</sup> likely originating from China. End-to-end email encryption is possible using the OpenPGP or S/MIME protocols but their use is not common, primarily due to their poor usability, which in turn is at least partially a result of having to stay compatible with older insecure email standards.

In contrast, instant messaging applications had more opportunity to start with a clean-slate (because there is no expectation of compatibility among different networks) and so this is where much innovation in terms of end-to-end security has taken place. Secure voice communication however has had less attention than instant messaging so in the remainder of the article we shall examine what should be expected of a secure voice communication system, and in particular see how one of the latest and up-coming protocols, MIKEY-SAKKE<sup>7</sup>, which comes with UK government backing, meets these criteria.

### *MIKEY-SAKKE and Secure Chorus*

MIKEY-SAKKE is the security protocol behind the Secure Chorus<sup>8</sup> voice (and also video) encryption standard, commissioned and designed by GCHQ through their information security arm, CESG. GCHQ have announced that they will only certify voice encryption products through their Commercial Product Assurance (CPA)<sup>9</sup> security evaluation scheme if the product implements MIKEY-SAKKE and Secure Chorus. As a result, MIKEY-SAKKE has a monopoly over the vast majority of classified UK government voice communication and so companies developing secure voice

<sup>2</sup> Vassilis Prevelakis and Diomidis Spinellis, “The Athens Affair”, IEEE Spectrum, 2007, <http://spectrum.ieee.org/telecom/security/the-athens-affair>

<sup>3</sup> Ryan Gallagher, “Operation Socialist”, The Intercept, 2014, <https://theintercept.com/2014/12/13/belgacom-hack-gchq-inside-story/>

<sup>4</sup> Jeremy Scahill and Josh Begley, “The Great SIM Heist”, The Intercept, 2015, <https://theintercept.com/2015/02/19/great-sim-heist/>

<sup>5</sup> James Bamford, “A Death in Athens”, The Intercept, 2015, <https://theintercept.com/2015/09/28/death-athens-rogue-nsa-operation/>

<sup>6</sup> David Drummond, “A new approach to China”, Google, 2010, <https://googleblog.blogspot.co.uk/2010/01/new-approach-to-china.html>

<sup>7</sup> Michael Groves, *MIKEY-SAKKE: Sakai-Kasahara Key Encryption in Multimedia Internet KEYing (MIKEY)*, RFC 6509, IETF, 2012

<sup>8</sup> CESG, “Secure Voice at OFFICIAL”, White paper, 2015, <http://www.cesg.gov.uk/guidance/secure-voice-official>

<sup>9</sup> CESG, “Secure Real-Time Communications Gateway”, CPA Security Characteristic, 2015, <https://www.cesg.gov.uk/documents/cpa-security-characteristic-secure-real-time-communications-gateway>

communication systems must implement it in order to gain access to this market. GCHQ can also set requirements of what products are used in the public sector and as well as for companies operating critical national infrastructure.

UK government standards are also influential in guiding purchase decisions outside of government and we are already seeing MIKEY-SAKKE marketed commercially as “government-grade security”<sup>10</sup> and capitalising on their approval for use in the UK government. For this reason, and also because GCHQ have provided implementers a free open source library<sup>11</sup> to make it easier and cheaper to deploy Secure Chorus, we can expect wide use MIKEY-SAKKE in industry and possibly among the public. It is therefore important to consider whether MIKEY-SAKKE is appropriate for wide-scale use. For the reasons outlined in the remainder of this article, the answer is no – MIKEY-SAKKE is designed to offer minimal security while allowing undetectable mass surveillance though key-escrow, not to provide effective security.

The EFF scorecard<sup>12</sup> gives a summary of some important security features for the diverse range of instant messaging applications and networks, so these serve as a useful starting point to develop security requirements of voice encryption.

<sup>10</sup> <http://www.armorcomms.com/solutions/>

<sup>11</sup> <https://bitbucket.org/securechorus/>

<sup>12</sup> <https://www.eff.org/secure-messaging-scorecard>

1. *Is your communication encrypted along all the links in the communication path?*

If the encryption is removed at any point, for example at the network provider for cellphone conversations, this creates a weak link and so should be avoided.

2. *Is your communication encrypted with a key the provider doesn't have access to?*

This criterion differentiates link encryption (which can meet the first criterion) from end-to-end encryption where only the communication partners have access to the unencrypted content. If this criterion is met, the communication can be secure even if the network provider's computers are compromised.

3. *Can you independently verify your correspondent's identity?*

It important to know who you are communicating with so as to prevent a “man in the middle” attack where the two partners think their communications are end-to-end encrypted when actually there is an eavesdropper who is removing the encryption, examining (or modifying) the unencrypted content then re-encrypting it before passing it on (see Figure 1). This criterion is only met if any such attack can be detected, even if the network provider's computers are compromised.

4. *Are past communications secure if your keys are stolen?*

Like the network provider, the computers used by the communication partners themselves are also subject to attack and the cryptographic keys they store may be compromised. This criterion states that should a key compromise occur, through use

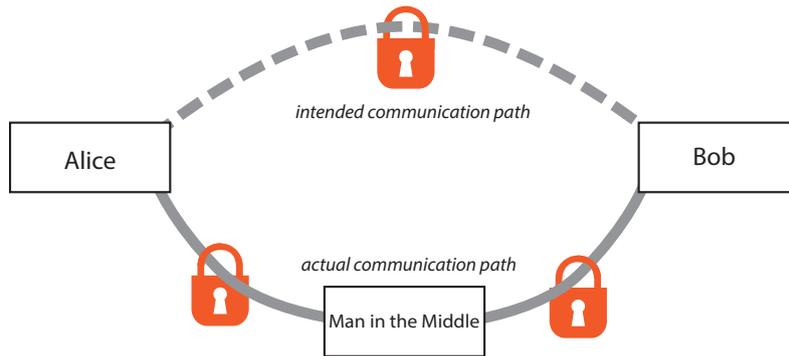


Figure 1: Alice thinks her communication with Bob is encrypted securely, but actually a man-in-the-middle is eavesdropping on the conversation.

of forward-secure cryptography, past communications must remain secure and only conversations started after the key was compromised will be vulnerable.

5. *Is the code open to independent review?*
6. *Is the crypto design well-documented?*
7. *Has there been an independent security audit?*

The final three criteria are on whether the implementation (5) and design (6) can be reviewed, and whether they actually have been reviewed (7).

Whether these seven criteria are the best ones to evaluate a product is an ongoing debate, but they are based on years of experience of using instant messaging tools in hostile environments and are representative of some of the ways by which communication security is breached in practice. In particular, a common thread throughout the criteria is a recognition that generally breaches occur as a result of flaws in the security protocol design, and particularly software implementations, rather than the underlying cryptographic algorithms.

Robust systems try to minimise the number of components which are in a position to break a user's security, such as the operator of the communication links (criteria 1 and 3) or the provider of the network (criterion 2). Robust systems should then try to build confidence in the security, of the design and implementation, of the remaining components which must be relied upon (criteria 5, 6 and 7). Finally, when compromises do happen the damage should be limited (criterion 4).

### *Key exchange for voice encryption*

To assess MIKEY-SAKKE against these and other criteria we need to explore more detail of how the security protocol works. MIKEY-SAKKE extends the Multimedia Internet KEYing (MIKEY)<sup>13</sup> standard, designed for voice and video encryption, by using Sakai-Kasahara Key Encryption (SAKKE)<sup>14</sup>. MIKEY just focusses on the

<sup>13</sup> Jari Arkko, et al., *MIKEY: Multimedia Internet KEYing*, RFC 3830, IETF, 2004

<sup>14</sup> Michael Groves, *Sakai-Kasahara Key Encryption (SAKKE)*, RFC 6508, IETF, 2012

most difficult part of secure communications – establishing a cryptographic key, using historically slow asymmetric cryptography – and leaves the actual job of encrypting communications with fast symmetric cryptography under the session key that was established, for implementers to select. The Secure Chorus standards however recommend the Advanced Encryption Standard (AES) in Galois Counter Mode (GCM) with a 128 bit key size, which is a widely used and efficient standard considered sufficiently secure for almost any purpose.

All the variants of MIKEY, and key-exchange algorithms in general, aim to ensure that at the completion of the protocol the legitimate communication partners share a session key that is not feasible for anyone else to infer. MIKEY supports the Ephemeral Diffie-Hellman (EDH) algorithm, which is widely used when both communication partners are online at the same time. The two communication partners each generate a random number using a cryptographic random number generator, and send this in encrypted form to each other. Then, using their own random number and the encrypted form of their partner's random number, each will be able to compute exactly the same session key. However, someone eavesdropping on the exchange will not be able to guess the key, even if they later compromise the computers of either or both of the communication partners.

EDH therefore allows the creation of systems secure against eavesdropping (meeting criteria 1 and 2), and offers forward security (criterion 4), but it is not itself resistant to man-in-the-middle attacks (criterion 3) because the attacker could perform a separate EDH exchange with each partner and so learn both keys. Therefore, EDH exchanges are usually digitally signed with a long-term asymmetric key to prevent man-in-the-middle attacks while still offering forward-security. However, to fully meet criterion 3 it is necessary to be able to securely verify this digital signature, which is discussed in more depth later.

Variants of EDH with digital-signatures underlie almost all end-to-end encrypted instant messaging systems, as by definition both communication partners are online during the conversation and so can agree on keys. The same situation applies to voice and video calls. However, for email and email-like systems where the recipient is not necessarily online, EDH cannot be straightforwardly used and so for OpenPGP and S/MIME the sender generates a session encryption key and encrypts it to the recipient's public key then sends this with the message encrypted under the session key. In this way end-to-end encryption is achieved. However, for this approach to be secure the sender must ensure they have the right public key for the recipient (and not that of a man-in-the-middle) and also because discovering the recipient's private key will allow all past messages to be decrypted, this approach does not offer forward secrecy.

## The design of MIKEY-SAKKE

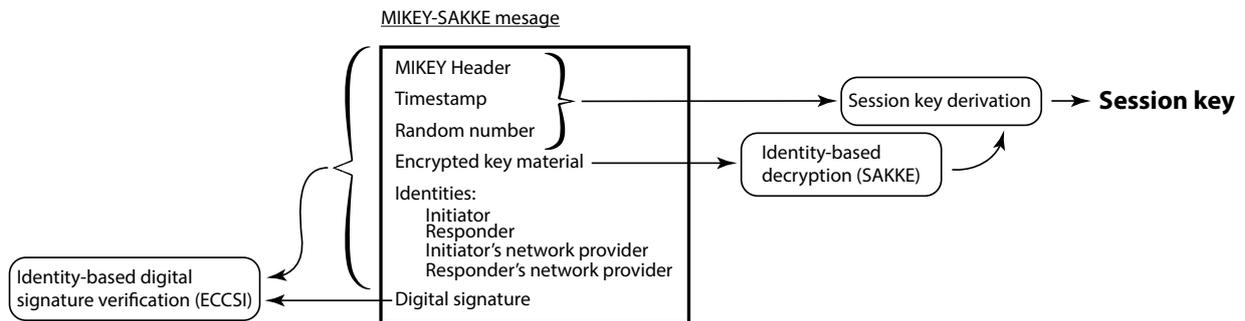


Figure 2: The MIKEY-SAKKE message is sent from the initiator to responder. The responder first checks the digital signature on the message (using the initiator's public key), then decrypts the key material (using the responder's private key). Finally, using the decrypted key material and other fields in the message, the responder can derive the session key.

MIKEY supports EDH but MIKEY-SAKKE works in a way much closer to email encryption. The initiator of a call generates key material, uses SAKKE to encrypt it to the other communication partner (responder), and sends this message to the responder during the set-up of the call – see Figure 2. However, SAKKE does not require that the initiator discover the responder's public key because it uses identity-based encryption (IBE). In conventional public key systems each party generates their own private key and distributes their public key to anyone who needs it but in an IBE system, all private keys are generated by the network provider from their master private key. The MIKEY-SAKKE message is also digitally signed to prevent tampering, using the Elliptic Curve-Based Certificateless Signatures for ID-based encryption (ECCSI)<sup>15</sup> algorithm – another IBE algorithm.

Using the responder's unique identity (e.g. email address or phone number), and the network provider's master public key, the initiator can compute the responder's public key and so encrypt the key material. To obtain the session key the responder needs to generate the initiator's public key from the network provider's public key, and before the call occurred have asked the network provider for the private key corresponding to their own identity. Using these keys, the responder can check the message for tampering then recover the key material, and finally derive the session key. While the network provider master key is valid for a long period, users' keys are valid for a month, so the network provider must keep the master key permanently available, to allow new users to join the network, and for existing users to periodically download the current private key that corresponds to their identity.

MIKEY-SAKKE has the advantage that only one key-exchange message is needed so call setup time is reduced compared to EDH. It also moves the complexity of key distribution: in standard public key systems the challenge is securely distributing public keys; in IBE the challenge is securely distributing private keys. However, IBE introduces fatal flaws for protocol security. Criterion 1 is met

<sup>15</sup> Michael Groves, *Elliptic Curve-Based Certificateless Signatures for Identity-Based Encryption (ECCSI)*, RFC 6507, IETF, 2012

(calls are encrypted from initiator to responder) but criterion 2 is not (the network provider generates the private key so can discover the session key and thus eavesdrop on calls). Criterion 3 fails because if the network provider is compromised then an impersonator could also know the responder's private key. Criterion 4 also is not met because past communications can be decrypted if the responder's private key, or network provider's master key, is discovered.

The existence of a master private key that can decrypt all calls past and present without detection, on a computer permanently available, creates a huge security risk, and an irresistible target for attackers. Also calls which cross different network providers (e.g. between different companies) would be decrypted at a gateway computer, creating another location where calls could be eavesdropped.

Criteria 5, 6 and 7 cannot be assessed because they apply to the product rather than the algorithms the product uses, though here MIKEY-SAKKE at least helps. The protocol is well documented as an Internet standard, is reasonably simple, and has been externally evaluated<sup>16</sup> to some extent. Other security protocols like TLS have complex option-negotiation steps which offer future-proofing but have probably been more a liability than asset<sup>17</sup>. MIKEY-SAKKE fixes most cryptographic parameters to reasonable defaults. Products which implement MIKEY-SAKKE may not be open source, but the GCHQ provided protocol implementation is. Products which go through CPA assessment will be audited by GCHQ, though the detailed audit report is not made publicly available.

### *The motivation behind MIKEY-SAKKE: key escrow*

The MIKEY-SAKKE approach is not the only possible IBE approach: equally IBE could digitally sign EDH key agreement messages. Such a design would meet Criteria 1 and 4, though still not 3. Criterion 2 would however be partially met because compromising the network provider would only allow active man-in-the-middle attacks and not passive eavesdropping, which substantially increases the difficulty of carrying out attacks. In fact such a protocol has been developed for voice encryption – Identity-Based Authenticated Key Exchange (IBAKE) Mode of Key Distribution in Multimedia Internet KEYing (MIKEY) – (MIKEY-IBAKE)<sup>18</sup>.

So this raises the question, why was MIKEY-SAKKE designed this way and why are GCHQ not permitting the use of EDH based voice encryption standards for UK government communications, and not supporting the deployment of such protocols elsewhere, despite their superior security? It certainly cannot be attributed to incompetence. GCHQ have extremely capable staff and the discovery that they had built the ability to eavesdrop on all Internet communications transiting the UK<sup>19</sup> demonstrates their technical skills. Therefore, there must be other design criteria behind MIKEY-

<sup>16</sup> Chloe Bell, *Analysing MIKEY-SAKKE: A Cryptographic Protocol for Secure Multimedia Services*, Master's thesis, Imperial College London, <http://pubs.doc.ic.ac.uk/mobius-mikey-sakke-analysis/mobius-mikey-sakke-analysis.pdf>, 2015

<sup>17</sup> Benjamin Beurdouche, et al., "A Messy State of the Union: Taming the Composite State Machines of TLS", in *IEEE Symposium on Security and Privacy*, 2015, pp. 535-552

<sup>18</sup> Violeta Cakulev and Ganapathy Sundaram, *MIKEY-IBAKE: Identity-Based Authenticated Key Exchange (IBAKE) Mode of Key Distribution in Multimedia Internet KEYing (MIKEY)*, RFC 6267, IETF, 2011

<sup>19</sup> Ewen MacAskill, et al., "GCHQ taps fibre-optic cables for secret access to world's communications", *The Guardian*, 2013, <http://www.theguardian.com/uk/2013/jun/21/gchq-cables-secret-world-communications-nsa>

SAKKE, or that the design follows culturally-held beliefs embedded in GCHQ about how protocols should work.

Although the words are never used in the specification, MIKEY-SAKKE supports key escrow. That is, if the network provider is served with a warrant or is hacked into it is possible to recover responder private keys and so decrypt past calls without the legitimate communication partners being able to detect this happening. Secure Chorus facilitates undetectable mass surveillance, in a way that EDH based key encryption schemes would not. This is presented as a feature rather than bug, with the motivating case in the GCHQ documentation being to allow companies to listen to their employees calls when investigating misconduct<sup>20</sup>, such as in the financial industry.

The aim of GCHQ's development of MIKEY-SAKKE – to weaken security of in order to facilitate surveillance – is made clear through their activity on the 3GPP standardisation committee responsible for “Lawful Interception (LI)”: ensuring that law enforcement and intelligence agencies are able to eavesdrop on 4G cellphone calls. The National Technical Assistance Centre (NTAC), the part of GCHQ responsible for assisting law enforcement and intelligence agencies with decryption and data analysis, sits on this committee (known as the “3GPP SA3 LI”) and their representative served as secretary.

GCHQ's submission to the committee<sup>21</sup> provides an analysis of MIKEY-IBAKE and points out the security features which result from its use of EDH – forward security and resistance to passive eavesdropping – are incompatible with their requirements of allowing large scale undetectable surveillance, both from a technical and legal perspective. Consequently GCHQ requested that MIKEY-IBAKE should not be used, stating:

“In light of these requirements, UK government has developed a similar scheme, MIKEY-SAKKE, which supports 3GPP SA3 LI requirements and has additional benefits such as low latency.”

In 2012, Alcatel-Lucent and Rogers Wireless proposed an alternative and more covert key-escrow approach to the same committee: deliberately weakening the random-number generation in the MIKEY-IBAKE EDH exchange<sup>22</sup> rather than replacing it with MIKEY-SAKKE.

Another way that MIKEY-SAKKE diverges from other secure instant messaging and phone systems is that there is no attempt to protect the identity of the communication partners, only the call content. In this way an eavesdropper can build up social network maps showing who is communicating with whom, when and how often, even if anonymising technology is used. This “metadata” is often more sensitive than the content – General Michael Hayden, former director of the US National Security Agency (NSA) stated “we kill people based on metadata”<sup>23</sup>. The design is not an accident – the GCHQ documentation states “MIKEY-SAKKE is an enterprise-level solution where anonymity is not possible”.

<sup>20</sup> CESC, “Using MIKEY-SAKKE: Building secure multimedia services”, White paper, 2014, <https://www.cesg.gov.uk/white-papers/using-mikey-sakke-building-secure-multimedia-services>

<sup>21</sup> NTAC, “LI of MIKEY-IBAKE, a UK perspective”, Report to 3GPP TSG-SA WG3-LI Meeting 38 (SA3LI10\_099), 2010, <https://cryptome.org/2014/03/nsa-uk-mikey-ibake.pdf>

<sup>22</sup> Christopher Parsons, “The Governance of Telecommunications Surveillance: How Opaque and Unaccountable Practices and Policies Threaten Canadians”, Telecom Transparency Project, 2015, <https://www.telecomtransparency.org/wp-content/uploads/2015/05/Governance-of-Telecommunications-Surveillance-Final.pdf>

<sup>23</sup> David Cole, “We Kill People Based on Metadata”, The New York Review of Books, 2014, <http://www.nybooks.com/daily/2014/05/10/we-kill-people-based-metadata/>

### *Implications of protocol design*

Key-escrow is an approach for building back-doors into encryption systems promoted by the NSA in the early 1990s and built into the Clipper encryption device<sup>24</sup>. With Clipper, a normal key exchange algorithm would be performed, but the resulting session key would also be encrypted under a separate escrow key held by a special department of the US government (the escrow agent). A third party wishing to listen to an eavesdropped encrypted call would request that the escrow agent decrypt the escrowed session key, and so allow the call to be decrypted. The escrow agent would only allow use of the escrow key after verifying legal authorisation and would keep records of such requests for oversight audits. After strong opposition from civil liberties groups, scientists and politicians the proposals were dropped by the NSA.

MIKEY-SAKKE allows third-party access to encrypted conversations in a different manner (sometimes its approach is called “key recovery” rather than “key escrow”). Firstly, the capability of a third party being able to decrypt past calls is integral to the MIKEY-SAKKE key-exchange process so is harder to bypass. One flaw in the Clipper protocol was that someone could use the chip to encrypt calls but prevent the escrowed key from being usable<sup>25</sup>, and such behaviour would not be possible to detect unless the escrow facility were actually used. Secondly, the control over access to recovered keys is different. In Clipper, the ability to decrypt escrowed keys would be with a government agency, possibly under joint control, and any use would be audited. With MIKEY-SAKKE, access to private keys would be provided by companies operating communication networks, and so may be more vulnerable to hacking, intimidation of employees or insider abuse, as well as allowing less oversight.

MIKEY-SAKKE is not the first attempt by GCHQ to promote a key exchange protocol that facilitates key escrow – the 1996 GCHQ protocol<sup>26</sup> differs in the detail (particularly as it pre-dates the SAKKE encryption algorithm by four years) but is similar in its characteristics. Like MIKEY-SAKKE, the GCHQ protocol is based on IBE, and private keys are distributed by a central authority who can then also decrypt eavesdropped messages. A notable difference however is that while the GCHQ protocol was explicitly stated to support key escrow to facilitate law enforcement and intelligence agency access, this controversial aspect has not been included in the description of MIKEY-SAKKE and instead the efficiency over EDH is emphasised.

This recurrence of key-escrow proposals from GCHQ is not surprising, given the conflict of interest inherent in making one agency responsible for both spying on communications and preventing spying. GCHQ designs the encryption technology used by government to prevent unauthorised parties having access to classified information. But GCHQ also wants the ability to examine how

<sup>24</sup> <https://www.epic.org/crypto/clipper/>



Figure 3: The AT&T TSD-3600E: one voice encryption product based around the Clipper chip. (Photo: Matt Blaze)

<sup>25</sup> Matt Blaze, “Protocol Failure in the Escrowed Encryption Standard”, in *ACM Conference on Computer and Communications Security (CCS)*, 1994, pp. 59–67

<sup>26</sup> Ross Anderson and Michael Roe, “The GCHQ Protocol and its Problems”, in *International Conference on the Theory and Application of Cryptographic Techniques (EUROCRYPT)*, 1997, pp. 134–148

this encryption technology is used to investigate suspected leaks whether to companies, the press, or foreign intelligence agencies.

Where the situation becomes more complex is when encryption technology used by companies is deliberately weakened in order to facilitate surveillance, and these weaknesses are then exploited by others. One such case is the recent news that the Dual-EC DRBG cryptographic random number generator in Juniper's network equipment, almost certainly designed as a back-door for the NSA<sup>27</sup> (which Juniper tried to close) but was later modified to give access to someone else<sup>28</sup>. As the Juniper case and the development of MIKEY-SAKKE shows, the increasing use of commercial software for securing government communications, and government-supported security software for securing commercial communications, makes it difficult to predict the wide-scale implications of design changes promoted by intelligence agencies.

Protocols like MIKEY-SAKKE also raise broader questions. Phillip Rogaway argues that cryptography is not politically neutral<sup>29</sup>, and so has moral dimensions as well as presenting intellectually stimulating puzzles for mathematicians. Cryptography has the capability of re-arranging power, and certain designs have fundamentally different characteristics. On Identity Based Encryption, Rogaway notes "one can easily see the authoritarian tendency built into IBE".

In fact, the implications of having centralised authorities who have the ability to break communication users' security goes further than just raising questions about privacy of communications. In "Do Artifacts have Politics"<sup>30</sup>, Langdon Winner shows that the design of some physical artifacts either requires or at least encourages particular social structures of power and authority. One example he gives is civilian nuclear power: because of the incredible damage which could come from a terrorist obtaining just a tiny amount of nuclear material, it would be considered necessary to put in place widespread surveillance and other intrusions on civil liberties to impose strict safeguards. In this way a technical artefact such as nuclear power can have deep and unavoidable political ramifications.

Building inherently fragile communications security systems also naturally leads to certain political structures. Where a compromised network provider would give an attacker the ability to undetectably decrypt any message both past and future, the justification for extraordinary effort (and budget) to protect these systems naturally follows. Not only must there be strong assurance that the critical aspects of the network provider work correctly, but also that network surveillance be put in place to detect and prevent such attacks. Background checks on employees, and strict secrecy over how security mechanisms work also seem likely. These impositions on civil liberties will have further implications than just the protection of the systems they were intended for.

<sup>27</sup> Daniel J. Bernstein, et al., *Dual EC: A Standardized Back Door*, Report 2015/767, Cryptology ePrint Archive, <https://eprint.iacr.org/2015/767>, 2015

<sup>28</sup> Matthew Green, "On the Juniper backdoor", blog post, 2015, <http://blog.cryptographyengineering.com/2015/12/on-juniper-backdoor.html>

<sup>29</sup> Phillip Rogaway, "The Moral Character of Cryptographic Work", Essay to accompany 2015 IACR Distinguished Lecture, 2015, <http://web.cs.ucdavis.edu/~rogaway/papers/moral-fn.pdf>

<sup>30</sup> Langdon Winner, "Do Artifacts Have Politics?", *Daedalus*, 109, 1980, pp. 121-136

*Towards secure voice communications*

Robust communication security systems avoid the single-point of failure of having centralised weak-points, such as the network provider in MIKEY-SAKKE-based encryption systems. EDH key-agreement gives most of the properties required, and is implemented for voice communications in the Secure Communications Interoperability Protocol (SCIP)<sup>31</sup> and Z Real-time Transport Protocol (ZRTP)<sup>32</sup>, among others. SCIP was developed in the NSA so is mainly used for government applications and ZRTP is mainly in civilian applications. The MIKEY-SAKKE design documentation explicitly states that these protocols were not able to meet GCHQ’s “scale and usability requirements”<sup>33</sup>, but does not expand on this claim.

One voice encryption application which ticks all the boxes is Signal<sup>34</sup> (formerly RedPhone) from Open Whisper Systems. Using ZRTP means that it is end-to-end encrypted (criteria 1 and 2) and offers forward security (criterion 4) – see Table 1. Its design and code are available for audit (criteria 5 and 6), and it has fared well to examination<sup>35</sup> (criterion 7). Resisting man-in-the-middle attacks (criterion 3) is the most challenging requirement as it either forces the user to rely on the security of third parties or carry out their own checks. For this reason, the major differentiating factor between voice encryption schemes is how these checks are performed.

<sup>31</sup> <https://www.iad.gov/SecurePhone/>

<sup>32</sup> Philip Zimmermann, et al., *ZRTP: Media Path Key Agreement for Unicast Secure RTP*, RFC 6189, IETF, 2011

<sup>33</sup> CESC, “Secure Voice at OFFICIAL”, White paper, 2015, <http://www.cesg.gov.uk/guidance/secure-voice-official>

<sup>34</sup> <https://whispersystems.org/>

<sup>35</sup> Matthew Green, “Here come the encryption apps!”, blog post, 2013, <http://blog.cryptographyengineering.com/2013/03/here-come-encryption-apps.html>

	1) Encrypted	2) Provider-safe	3) Verified identity	4) Forward-secure
MIKEY-SAKKE	✓	✗	✗	✗
MIKEY-IBAKE	✓	partial	✗	✓
ZRTP	✓	✓	✓	✓

Table 1: Comparison of three protocols against the EFF criteria.

Signal, in common with other ZRTP-based designs, uses Short Authentication Strings where not only does the EDH key-agreement result in a session key but also a 16 bit value which is converted into two words which are shown on the screen. One caller must read out the words, and if there is no man-in-the-middle, there is only one EDH exchange, so the other caller will see that they match. However this is not foolproof: the callers must follow the procedure exactly and even then it’s vulnerable to an attacker impersonating the voice of the other caller<sup>36</sup>.

Signal’s instant message encryption facility (formerly TextSecure) offers another approach to preventing man-in-the-middle attacks:

<sup>36</sup> Maliheh Shirvanian and Nitesh Saxena, “Wiretapping via Mimicry: Short Voice Imitation Man-in-the-Middle Attacks on Crypto Phones”, in *ACM Conference on Computer and Communications Security (CCS)*, 2014, pp. 868–879

allow callers to meet in person before calling, and verifying that the “fingerprint” of the other caller’s long-term public key matches what they expect. This same approach could be applied to voice encryption too. Moreover, Signal’s instant message facility uses a variant of EDH – Axolotl – which offers forward security even when one party is offline. This facility could be used for end-to-end encrypted voicemail, which is one of GCHQ’s motivating cases for why to use MIKEY-SAKKE rather than EDH in standard MIKEY.

While cryptographically sound, experience with other applications has shown that manual fingerprint verification is hard for non-experts to perform and as a result has been hidden in a part of the Signal user-interface where only determined experts will find it. This is an example of another aspect where the authors of Signal have put great effort: making difficult choices to help make the software easy to use and so have widespread use.

Although not one of the EFF criteria, encouraging widespread use of cryptographic tools is also a security advantage. If encrypted messaging is rare, those who use it could be singled out for attack, whether harassment of the person or hacking of their computer. Good usability is necessary to achieve this goal: if secure calls are much harder to make than insecure ones, few people will bother and those who do may expend so much effort struggling with the software that they will not have enough attention remaining to detect man-in-the-middle attacks.

Also necessary for widespread use is the ability to discover which contacts support which secure voice standard, and obtaining the contacts’ public keys. Signal uses people’s phone numbers as an identifier and has a central service which allows people to discover which people in their phone’s contact database also have Signal. This centralisation is great for usability but does put Signal users’ metadata somewhat at risk<sup>37</sup>. Meta-data protection, both for contact discovery and calls, in large-scale networks is a difficult and as yet unsolved problem.

<sup>37</sup> Moxie Marlinspike, “The Difficulty Of Private Contact Discovery”, blog post, 2014, <https://whispersystems.org/blog/contact-discovery/>

### *Preventing Man-in-the-Middle attacks*

There are other options available for man-in-the-middle resistance. Key-continuity, also known as Trust On First Use (TOFU) relies on the fact that while man-in-the-middle attacks are possible they are hard to perform consistently as a user moves between different networks. In TOFU-based schemes the first public key used by a contact is stored, and any change of this key is flagged up to the user as being cause for suspicion that a man-in-the-middle attack has started (or stopped). Secure Shell (SSH) popularised this model and another ZRTP implementation – Silent Phone<sup>38</sup> – uses this in addition to Short Authentication Strings.

<sup>38</sup> <https://www.silentcircle.com/>

For HTTPS encrypted web browsing the Certification Authorities (CA) system is intended to prevent man-in-the-middle attacks. The same system is used for S/MIME encrypted email and could

equally be applied to encrypted phone calls. CAs are organisations which certify that a particular public key corresponds to a particular name (e.g. domain name or email address). This approach is convenient for the users because their web browser does all the checking for them, but a malicious or compromised CA can harm users so they become tempting targets for attack. In 2011 two certification authorities were compromised<sup>39</sup> by hackers affiliated with or supporting the Iranian government. Preventing such attacks is difficult but Certificate Transparency<sup>40</sup> aims to help detect such attacks quickly to allow better mitigation.

OpenPGP encrypted email implements an extension of key fingerprint verification, called the Web of Trust where users not only check their contacts' key fingerprints but can also choose to rely on checks that those contacts perform on their contacts, and so on. Because who is a contact of whom becomes publicly visible, this is a disaster for metadata protection, and the OpenPGP web of trust database is now built into tools for covert intelligence gathering<sup>41</sup>. For this reason and for usability reasons, I am not optimistic about using the Web of Trust for secure phone calls and even for email the leading OpenPGP implementation, GnuPG now implements TOFU<sup>42</sup>.

A final method for man-in-the-middle protection is used by the OTR (Off the Record) instant messaging protocol<sup>43</sup>, but could equally be applied to secure voice communication. Here the communication partners agree on a short password (perhaps by meeting in person), or work out a question to which only the legitimate partner will know the correct answer. This password or question-and-answer is combined with the EDH key exchange, such that it is not possible for an attacker to perform a man-in-the-middle<sup>44</sup> without knowing the secret information.

### *Supporting investigation of misconduct*

None of the ZRTP-based implementations I've mentioned supports key escrow, but for some niche applications third-party access to encrypted calls may be necessary. The MIKEY-SAKKE documentation suggests the regulated financial industry is one such case, but in reality what they need is quite different from what MIKEY-SAKKE offers. MIKEY-SAKKE means that encrypted calls that are recorded can be decrypted indefinitely into the future, because the network provider has a long term private key from which all user keys can be generated. This is not what the financial industry wants, firstly because they require not only that recorded calls are kept for the legally mandated time, but also that they are permanently deleted immediately after this period. Secondly, key escrow is only useful if the encrypted calls are recorded as a matter of course, and financial companies don't do this.

For these reasons it is better to build regulatory-mandated call-recording systems on top of secure phone systems and record the

<sup>39</sup> Eva Galperin, et al., "A Post Mortem on the Iranian DigiNotar Attack", blog post, 2011, <https://www.eff.org/deeplinks/2011/09/post-mortem-iranian-diginotar-attack>

<sup>40</sup> <http://www.certificate-transparency.org/>

<sup>41</sup> <https://www.paterva.com/>

<sup>42</sup> Neal H. Walfield, "TOFU for GnuPG", Email, 2015, <https://lists.gnupg.org/pipermail/gnupg-users/2015-October/054608.html>

<sup>43</sup> Nikita Borisov, et al., "Off-the-Record Communication, or, Why Not To Use PGP", in *ACM Workshop on Privacy in Electronic Society (WPES)*, 2004, pp. 77–84

<sup>44</sup> Chris Alexander and Ian Goldberg, "Improved User Authentication in Off-the-Record Messaging", in *ACM Workshop on Privacy in Electronic Society (WPES)*, 2007, pp. 41–47

calls before they are encrypted, or after they are decrypted on the other side. Indeed, this is how today's products for call-recording in the financial industry work<sup>45</sup> and there's no reason to change.

Financial companies don't have perfect security, so having a server permanently available storing the master private key for all their communications is a huge risk, but unavoidable with MIKEY-SAKKE. If call-recording uses separate mechanisms to end-to-end encryption the call-recording key can be kept offline and only used in the exceptional circumstances of investigating suspected misconduct. In this way recordings could also be deleted when there is no regulatory or business reason to keep them, and at that point anyone who has eavesdropped on the encrypted call would not be able to decrypt them either (due to forward secrecy). It is also advisable for the financial industry to use the same applications as other people, and use products that protect metadata, because being singled out as a rich banker puts staff at risk if they are in countries where kidnapping is a problem.

<sup>45</sup> "Vodafone Mobile Voice Recording", <http://www.vodafone.com/business/global-enterprise/enterprise-managed-mobility/mobile-voice-recording>

### *Conclusions and future work*

The design of MIKEY-SAKKE is motivated by the desire to allow undetectable and un-auditable mass surveillance, which may be a requirement in exceptional scenarios such as within government departments processing classified information. However, in the vast majority of cases the properties that MIKEY-SAKKE offers are actively harmful for security. It creates a vulnerable single point of failure, which would require huge effort, skill and cost to secure – requiring resource beyond the capability of most companies. Better options for voice encryption exist today, though they are not perfect either. In particular, more work is needed on providing scalable and usable protection against man-in-the-middle attacks, and protection of metadata for contact discovery and calls. More broadly, designers of protocols and systems need to appreciate the ethical consequences of their actions in terms of the political and power structures which naturally follow from their use. MIKEY-SAKKE is the latest example to raise questions over the policy of many governments, including the UK, to put intelligence agencies in charge of protecting companies and individuals from spying, given the conflict of interest it creates.

*An edited version of this article appears in the March 2016 special edition of IEEE Computer Magazine: Communications and Privacy under Surveillance (S. J. Murdoch, "Insecure by Design: Protocols for Encrypted Phone Calls," in Computer, vol. 49, no. 3, pp. 25–33, Mar. 2016. doi:10.1109/MC.2016.70).*

 © 2016 Steven J. Murdoch. This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License. Typeset using Tufte- $\LaTeX$ .