Disruptive Analysis

Don't Assume

Data over Sound Technology:
Device-to-device communications & pairing without wireless radio networks

A Disruptive Analysis *thought-leadership* paper

June 2017

Author: Dean Bubley
Contact: information@disruptive-analysis.com

Introduction

We are starting to see the early signs of a surge in wireless Data-over-Sound (DoS) applications: the use of audio signals through the air, as a mechanism for sending data between devices. While the concept is not new, its current growth is being fuelled by the rise of smartphones, IoT devices and improved signal-processing technology. The need for simple device-to-device communications, transcending the limits of radio-based wireless systems like Bluetooth or NFC (Near Field Communications), is likely to allow DoS to carve some important niches.

DoS will mostly be used for indoor, low data-rate "proximity" applications, where ease-of-use requirements preclude multi-step "pairing" between end-points or point-to-multipoint uses. Even though only small amounts of data can be sent via DoS, this can include web addresses or triggers for other connected apps, which can then access richer content or data resources, via a normal wireless or wired connection.

Applications will be very diverse – from connecting children's toys, to industrial robots, to businesses' conference-room systems. When one considers the huge range of products equipped with microphones and speakers – as well as the possibility to extend range by sending audio signals via the phone network or public-address systems – the scope of DoS will be limited only by developers' imaginations.

There will be an important divide between audible and non-audible (essentially ultrasound, or "hidden") DoS applications. Important questions of trust and privacy will arise around the latter – there will need to be clear pre-existing consent if devices are connecting without obvious notification signs for their owners.

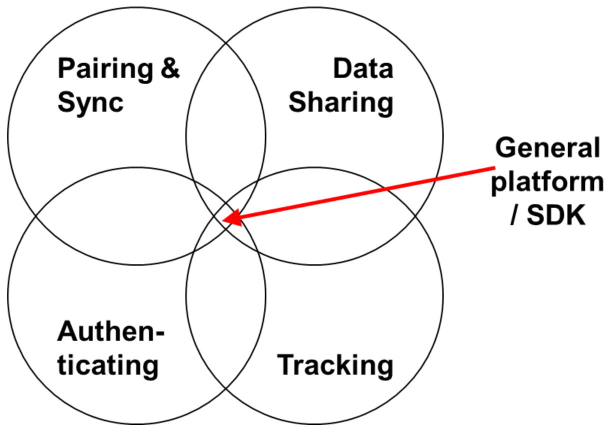
The value chain for DoS solutions is still emerging. We will likely see a mix of dedicated "vertical" integrated propositions for particular applications, alongside more-extensible "horizontal" APIs or PaaS models that can be used by general developers and systems providers. In some instances, the capability may be built directly into chips, OSs or devices, while in other cases it will be implemented in-app. We can also expect to see hybrids with other connectivity and signalling platforms, for example combining data-over-sound with short-range radio technologies, or adding DoS capabilities to broader voice/audio API platforms.

Some big names such as Google and Cisco already have point-product DoS integrations, and verticals such as mobile payments have become "real" with startups such as ToneTag gaining traction. Generalist technology providers such as this report's sponsor Chirp are becoming more prominent across multiple application domains.

As with all new technologies, there is a measure of early hype around DoS – and some overly "ambitious" claims by vendors trying to garner attention. Disruptive Analysis recommends developers and investors carefully evaluate what is real and proven, what can be reasonably expected in future, and what is just hot air.

Overall, Disruptive Analysis believes that DoS technology will follow a similar evolution path to other new app/device-capabilities. An analogy can be made with the rise of WebRTC, which embeds realtime voice/video streams into contextual communications. It requires specialist skills at first, even as theoretical costs come down. As awareness and expertise rises, an ecosystem of support functions emerges to allow the technology to be fully "democratised". DoS will eventually acquire testing suites, developer tools, platform support, standards and open-source elements on its way to become truly mainstream.

Data-over-Sound spans multiple application verticals, plus general platforms



Source: Disruptive Analysis

Structure & objectives of the paper

This document is structured into the following sections:

- Why is Data-over-Sound useful?
- Key technology enablers of DoS
- Example use-cases and applications for DoS
- Commercialisation pathways, opportunities & risks
- Conclusions & recommendations

The document has been prepared by independent research firm Disruptive Analysis, and commissioned by Chirp Ltd., for distribution to its customers, partners and a wider audience. It is based on Disruptive Analysis' continuous research programme covering IoT networking, service-provider and enterprise communications, and the "future of voice & messaging".

It should be read by CIOs, strategy executives, CTOs, CMOs, enterprise architects & planning/operational staff at major enterprises, communications service providers, information providers, software vendors, IoT firms, cable operators, ISPs, integrators, developers, XaaS providers and similar organisations.

Mentions of companies and products in this document are intended as illustrations of market evolution and are not intended as endorsements or product/service recommendations. For more details, please contact information@disruptive-analysis.com

Why is Data-over-Sound useful?

Data-over-sound technology encodes information into audio format, either audibly as "bleeps and tones", inaudibly above human hearing range, or "hidden" as imperceptible modifications of existing speech or music. It is received through a microphone, and decoded on another device. Depending on the implementation, closeness of the devices and environment, it can operate at data-rates of around 50-100bit/s, usually within a room.

DoS implementations can be divided between:

- Audible frequencies (generally in the range 30Hz-18,000Hz = 18kHz)
- Inaudible frequencies (for most purposes here, 18-20kHz, although true ultrasound frequencies go much higher)

It is reasonable to question why a relatively slow, and seemingly-primitive form of communication like DoS is needed today at all. We are constantly surrounded by multi-megabit (and soon gigabit) per-second wireless transmissions of data using WiFi, 4G and various other technologies. We connect our phones to headsets or vehicles with Bluetooth, or conduct transactions with QR-codes or NFC short-range wireless. Trade shows and industry news are filled with early prototypes of 5G technology, or low-power, wide-area IoT-centric systems like LoRAWAN and NB-IoT.

Yet despite these innovations, there are still various reasons why they cannot cover all eventualities and use-cases. Sound-based data transmission has a range of important characteristics and strengths that cannot easily (or reliably) be replicated with other mechanisms.

A Rich History

The use of sound waves, or audio-based electronic systems to send data (as well as speech) is not new. Humans have used drums and other percussion instruments to send messages, as well as create music, for thousands of years. Other animals use audio waves as well – such as bats' ultrasound or dolphins' sonar for echo-location – as well as more-normal communications between species' members.

For decades, telegraph and telecom networks have used sound technology as well. Some has been "electronic sound transmission" such as the familiar touch-tone (DTMF, dual-tone multi-frequency signalling) controls on handsets, or modems' encoding of data onto phone lines, the latter invented by IBM in the 1940s. Others have transmitted data in acoustic form, through the air – or even through water or solid objects.

Numerous research papers and niche examples have examined the data-over-sound concept for years, such as one published in IBM's Systems Journal in the year 2000[1]. Nevertheless, despite broad academic awareness and numerous proofs-of-concept, real-world applications of this type of approach have grown only slowly.

[1] http://ieeexplore.ieee.org/document/5387014/

However, the last few years have shown a sudden new explosion of interest in audio-based signalling, sensing and data transmission – alongside the growth of more familiar RF-type wireless data methods. Numerous DoS startups – with varying levels of vision and credibility - have emerged, as well as interest from established companies such as Cisco and Microsoft.

The advent of smartphones (often with higher-quality audio components and improved signal-processing capability) has meant that applications can take advantage of both audible and inaudible tones – aided by the third-parties offering "embeddable" functions via developer tools and APIs. The rise of IoT devices is adding a further huge impetus.

As this document will describe, sometimes the audio approach to data communications has better applicability than even such successful technologies as a Bluetooth. Both audible and non-audible use cases are now being commercialised, ranging from in-store retail/advertising "beacon" applications, through to signalling between industrial robots.

Quiet RF environments

Perhaps the easiest and most-understandable reason to use sound is that some locations *cannot* use radio. For reasons of safety or interference with other systems, various locations are "quiet". While the rules on using wireless devices on aircraft are being relaxed (eg for in-flight WiFi), there are still other contexts where there are still strict prohibitions or practical constraints:

- In hospitals, especially intensive-care units or near X-ray machines or other scanners
- Industrial facilities where welding or other equipment can blast out huge amounts of RF interference
- Mining, military and oil/gas installations which store and use explosives on-site, and which often prohibit RF use
- Sensitive facilities such as military or nuclear sites, where radios are strictly limited
- Scientific laboratories with delicate instruments, or indeed conducting research into RF phenomena themselves
- Buildings where wire meshes or metalwork can act as RF screens or filters

In such locations, the ability to transmit data via audio signals has huge potential. While some applications can obviously use wired or fibre connections, this is not possible where endpoints are mobile.

Usability and ubiquity vs. alternative wireless options

Many applications, especially on smartphones, need "proximity" connectivity. This can either be between two people and their phones, or between a phone and a location-specific system like a point-of-sale terminal or beacon.

There are multiple ways for short-range connections to be set up to transfer data – Bluetooth (including the BLE low-energy variant), NFC contactless chips, QR-codes displayed on screens, WiFi in P2P mode, proprietary methods like Apple AirPlay or AirDrop, and so on.

While these approaches all have their uses, there is often either "friction" in setting them up initially, or a problem with accessibility for *all* relevant users. A significant proportion of people have Bluetooth switched off, for example – or do not have NFC present or configured on their smartphones. Bringing devices close enough to read contactless chips or QR codes is not always possible. Even where there is a theoretical match in device capabilities, there may be too many steps involved in "pairing", or resistance by some users to doing something perceived as complex or "geeky". In environments where there is nobody to give advice or basic technical support – even just navigating menus – that is a major impediment. Often, users will give up on an application if it doesn't work easily and intuitively the first time.

In such instances, using sound as a primary (or secondary) approach to data-transfer may work better. More devices have microphones and speakers, compared to different wireless options, or cameras. There is less dependency on operating systems or OS version. Devices can be distances of a few metres or more apart. Legacy products can be supported more easily. While certain applications like the telephone dialler can be set to "mute", it is rare for the overall microphone capability to be switched off. Ultrasound or near-ultrasound can be used without intrusion or audible interruption – although this needs to be done with user consent and awareness.

One further advantage with data-over-sound compared with NFC is that it is full-duplex: on smartphones and most other devices, the microphone and speaker can send and receive data simultaneously.

If DoS approaches can improve the first-time user experience of interconnecting an application, for a greater number of people/devices, there is significant upside for the developers in building longer-term loyalty and usage.

Hybrids are also inevitable. Some vendors blend Bluetooth with audio-based signalling. Signal360, for example, creates retail/advertising beacons which support both technologies. It cites an estimate of 35% of phones which have Bluetooth switched off and which are therefore unreachable with that technology. Google's Nearby messaging function is primarily Bluetooth-centric, with ultrasound as a (poorly documented) alternative path. Other combinations of DoS + alternative communications mechanisms are likely to emerge, reflecting user-preferences and environmental constraints (audible DoS is not likely to be popular for ordering interval drinks in a theatre).

One other channel being explored for through-the-air data transmission is the use of light signals. Although visible light (or near-visible infrared or ultraviolet) waves are part of the same electromagnetic spectrum as RF, they have very different characteristics, as they have frequencies hundreds of thousands of times higher than most cellular or short-range radios systems (430–770 THz, compared to 700-2600MHz for most 3G, 4G, WiFi and Bluetooth). Although there are some interesting use-cases for so-called "Li-Fi" such as broadcasting data via LED lighting, there is very limited overlap with data-over-sound, especially given the reliance on line-of-sight communications and devices with suitable camera/sensor support.

IoT and legacy support

A critical benefit of DoS is that its use can be extended more-easily to a variety of existing devices and systems than other signalling/data-transmission methods. Far more objects have microphone and/or speaker support already, compared to either camera/screen or radio systems such as Bluetooth.

Door entry-phones, music systems, industrial machines, point-of-sale terminals, walkie-talkies, street-furniture, vehicles, old desk-phones and many other devices all have speakers and often microphones as well. While they may not have sophisticated computing capabilities or audio-processing, they may still be "hackable" by developers, to bridge them with other devices that can offer those functions. For example, a bank ATM machine, installed in a temporary location, could prove its authenticity (and transmit virtual funds) to a banking app or "wallet" on a user's phone, via DoS.

Physical/environmental considerations

Audible and near-audible sound waves have a number of important characteristics, that can both make DoS especially useful, or preclude it. Among the key considerations are:

- Limited transmission through building materials and walls. This means that sound is (mostly) confined to one room in a building, assuming doors and windows are shut and the volume levels are limited. This contrasts with various RF forms such as WiFi, which can "leak" much more easily. Indeed, the two could be used together – for example a café's guest WiFi being restricted in use to those people physically inside it. Sound can, however, travel around corners – and is less-prone to disruption from unexpected phenomena like rain or slow-moving objects.
- Slow speed of sound waves. In air, sound travels at about 300m/sec. This means that echoes as the sound bounces off walls can interfere with each other, and also take a significant amount of time to die down. This is one of the factors limiting the data rate of the technology.
- In some circumstances, ultrasound signals may have a longer latency than audible ones. This is because acoustic attenuation and other effects may influence higher-frequency sounds more than lower-band signals, meaning that they need to be repeated in a loop several times, for reliable reception by the microphone. Also, ultrasound data rates are lower, as they can only use about 2kHz (between about 18-20kHz), rather than the effective 15kHz available in audible range.
- Related to echoes, the slow speed of sound in air also results in the Doppler Effect, where frequencies of perceived sound become noticeably higher or lower, for fast-moving objects moving towards or away from the microphone. (An example is the changing engine tone of a passing car). While this presents additional processing challenges, it also allows motion-detection along with data transmission. This effect is proportional to frequency, which may mean that higher-frequency audible or ultrasound data becomes corrupted or harder to receive.
- Sound transmission can be affected by local furnishings and decorations in a room, or specifically-designed "deadening" materials.

Key DoS technology enablers & components

While a lot of researchers and companies have done work on data-over-sound technologies over the last 20+ years, not all have been able to create effective and commercialisation-ready solutions. While some of the lack of historic success can be attributed to poor business models, or insufficiently-sized user base, there have also been a variety of technical limitations.

It is important to consider what has changed – and also the sources of differentiation for solution providers. This section gives an overview of some of the key terminology and system components of DoS. (It does not however give detailed descriptions of the underlying science and information-theory, which are outside the scope of this document, but widely available from online sources).

Watermarking vs. Modulating

There are two basic ways to encode data into sound waves:

- **Modulating:** This method is a standalone form of creating DoS signals. It takes an existing source of data, and translates it into tones or pulses, based on a variety of possible coding schemes. As it does not require any other ongoing audio track, it tends to be more useful for ad-hoc and real-time communications between devices.
- **Watermarking (also steganography or audio-hiding):** This involves changing an existing audio signal such as music, by putting additional (inaudible) data into the stream. This is often used for purposes such as copyright protection, or putting extra meta-data into a particular source (eg artist or track-title). Because a host audio stream is required, this type of approach is often used for one-to-many or broadcast purposes, rather than ad-hoc device-to-device communication.

Acoustics and encoding methods

Acoustic data communications requires specific ways of setting up links and encoding information (protocols). In common with other forms of communications, there is a "stack" of different protocols that cover signalling, identity, error-correction, security/encryption and all the other "machinery" needed to get the overall system to function well. There are numerous trade-offs for speed vs. accuracy, and the impact of these balances varies depending whether sound is audible/inaudible, and on factors such as the relevant physical environment and the quality of the microphones and speakers used.

Although a lot of DoS implementations are superficially similar as they often use FSK (frequency-shift keying), the devil is in the detail. Getting acoustic data transmission perfected is *hard*. Every environment behaves differently – not just indoor and outdoor variations, but also effects such as absorption, reverberation, reflection and occlusion. We are all familiar with the impacts of different venues' designs on speaking or music performance – a similar set of considerations apply to data-over-sound as well.

Ideally, applications, devices and mics/speakers will be "tuned" to specific environments – think of it as analogous to a band's roadies and sound-engineers needing to reconfigure their audio setup for each location on a tour. A large factory with machinery will differ from an office boardroom, a forest, or a home with lots of soft furnishings. Resilience to real-world noise is important. There may be trade-offs between data rate and accuracy, or the use of customised frequencies or tone-lengths to best-avoid clashes with ambient noise.

There will be differences for applications where the devices (and their microphones / speakers) are well-known and provided by a single supplier, for example for robots in a factory. Conversely, if there has to be a "lowest common denominator", different priorities will be relevant – for example if an Android application needs to work both on a brand new Samsung Galaxy S8 and a 3-year old $50 device made by an 3rd-tier generic handset brand.

There are also other practical considerations at a device level – it is not always possible to control audio volume-levels, while over-driving speakers can lead to distortion or even damage. Some phones or tablets may be used in sound-absorbing cases, have dirt in speaker/mic openings and so forth.

Microphones & speakers

Most adults have a hearing range from around 20Hz (deep bass) up to around 16-17kHz, although children can sometimes hear up to 21-22kHz, and cats and dogs can hear up to 40kHz ultrasonics (hence inaudible dog-whistles). Most standard audio components in common personal devices work well up to around 20kHz, and some have usable ranges considerably higher. This means that use of quasi-ultrasound in the 18-19kHz range is feasible for many use-cases, notably those involving smartphones.

That said, not all legacy or larger-scale audio systems can handle those near-ultrasound frequency ranges. Some speakers such as those in cheaper phones, venue public-address (PA) systems, older TVs, door-entry systems and others products have more-limited frequency ranges. Other devices may have microphones or speakers inside an enclosure of some sort (for example, a soft toy) which limits their audio capabilities.

Integrations, APIs, SDKs & PaaS

Like most computing/signal-processing functions, DoS capabilities can either be developed independently by a device or solution vendor with sufficient expertise and resource, or obtained in software-component form from 3rd party suppliers. Over time, DoS might also become an underlying "native" capability of some device OSs or chipsets.

Where the technology is brought in from an external source, it will typically be made available either as "developer kits", or cloud-based platforms (or a combination of the two).

- Online / cloud-based APIs and SDKs tend to offer scalable capacity and easier integration with other functions – but require always/mostly-on connectivity from the device, and ongoing (ie scalable) cost commitments. They may either be

targeted to a specific vertical (for example, marketing-related, or for IoT connectivity) or could be broad horizontals usable for many applications.
- Offline / device-only SDKs are usually more limited in scale, but imply less reliance on both ubiquitous network connectivity and the originating provider's business model. However, they make it possible for developers to better tune the solution to the "affordances" of each specific platform or device.

Data-over-sound SDKs are available from various participants, for assorted platforms – most obviously iOS, Android and Windows, but also embedded platforms for IoT and even low-end computing environments such as the Raspberry Pi. The ability to create cross-platform functions is very important, as well as supporting legacy versions of OS and those running on both high- and low-end devices.

At the moment, the lack of DoS standards and multi-niche nature of the technology make it very hard to create truly "horizontal" platforms, especially those that look like familiar platform-as-a-service plays in other areas of technology such as Amazon AWS or Twilio. In addition, the risk of irresponsible use of some DoS technology (especially ultrasound) is quite high – it could harm small and growing players if their solutions were used in privacy-invasive applications, for example. As a result, we will probably need to wait for pure DoS-PaaS offers, where developers can simply sign up with an ID and credit card, or get unrestricted access to SDKs.

Integration with voice / audio systems

In some circumstances, it is possible to combine short-range data-over-sound technologies with other communications or audio systems. This can extend the range locally, provide broadcast capability, or allow hybrid services based around phone networks and enterprise UC platforms.

The main platforms to consider are:

- Public phone networks (also called PSTN)
- Business phone and conferencing systems (also called UC, unified communications systems)
- In-building PA (public address) systems.
- Broadcast TV and radio
- Web audio/video/game streaming
- Voice and video chat/conferencing features built-into websites and mobile apps

This can (in theory) allow sound-based information to be included in TV programming or advertising, or broadcast to anyone on a conference call, or throughout a large venue. However, there are numerous subtleties to consider here, especially if the originator of the DoS message has no control or visibility of the end users' devices, or acoustic environments.

In general, many existing audio and telecom systems are narrowband – intended just for a subset of audible frequencies. Even so-called "HD" voice systems are usually not designed to be capable of reaching above 15kHz – and normal telephone networks are often limited to 3kHz. In addition, microphones and speakers are of varying qualities and performance characteristics.

Building PA systems are also highly variable – although individual locations could be measured for particular applications. Broadcast systems are generally OK for audible frequencies but cannot cope with ultrasound. Enterprise conferencing is a bit more manageable, and sometimes has in-built mechanisms for reporting of quality

Many boxes in the middle of telecom networks (eg transcoders) strip out ultrasound frequencies, in order to compress traffic or translate between different audio encoding schemes (codecs). Streamed video / audio services such as Netflix and Spotify reduce the quality and amounts of data, based on the available network capacity and speed – these too, are unsuitable for ultrasound. In-building PA systems, door entry-phones, and many TVs are also unsuited to non-audible sounds.

The role of standards

At the moment, most DoS systems are highly proprietary, either customised for particular products, or provided through a specific vendor's SDKs. While that makes sense in the current phase of market evolution, Disruptive Analysis believes that standardisation will be needed, if DoS is to be truly "democratised" and used routinely in a similar fashion to Bluetooth or QR-codes for device-to-device communication. As well as technical standards allowing interoperability of different equipment, there may also need to be implementation guidelines or industry codes-of-conduct, for example around privacy and consent.

Given the broad range of use-cases for DoS, it is unclear exactly which standards body or bodies are best-placed to represent the interests of the sector. It needs to span browsers, mobile apps and embedded-computing devices for IoT, as well as consumer, business and industrial use-cases. For instance, the IEC (International Electrotechnical Commission) defines some electroacoustic standards for microphones and speakers – but we could also envisage the W3C (Worldwide Web Consortium) looking to improve its WebAudio or other browser APIs to reflect the requirements of ultrasound transmission. Individual standards applied to usage in specific industrial sectors (eg medical equipment and hospitals) are also going to be relevant.

It is likely that various organisations will ultimately become involved – some regarding technical performance, some around implementation and APIs, and some around security and ethics codes. We may also see regulatory intervention in some domains, for example as an extension of privacy legislation.

Analytics and AI

Over time, we can expect DoS systems to become much more self-configuring and self-healing. They may use additional sensors, software tools or web-based services, to deduce the type of environment they find themselves in. The software itself will be able to understand the limitations of the device it is installed on, or that of peers with which it is communicating.

Machine-learning may help underlying patterns to be detected, which help improve error rates, or which minimise interference to end-users. It is likely that such techniques will also be used to work out whether to use DoS at all, or if particular individual users / environments show preferences for Bluetooth.

Use-cases

This section of the document gives examples of applications for data-over-sound technology. Some of these are existing, real-world implementations, while others are more forward-looking or theoretical. While a number of important categories such as mobile payments, consumer-electronics pairing and retail/marketing are described in some depth, readers should be aware that there is a very "long tail" of possible use-cases in all manner of additional niches.

The use-case sections also give examples of companies active in each area. However, readers should be aware of wide variations in maturity – some are engaged in full-scale commercial deployments, others just in prototype / proof-of-concept phases – and some merely state hypothetical future aspirations. Mentions of companies are illustrations, not endorsements – and developers and other parties should carefully evaluate claims.

It is also important to recognise that the marketplace is highly dynamic – and the same outcomes may be achievable through various other approaches such as Bluetooth or QR-codes. Given that both the alternative technologies' capabilities, and the installed base of devices (eg smartphones or consumer IoT) are rapidly evolving, it does not automatically follow that *possible* DoS applications will always be commercially successful or adopted by end-users.

Before diving into the various sub-categories of application, it is also worth putting "Data-over-Sound" into context more fully. There are various top-level ways in which audio waves – audible or inaudible - can be used for connectivity and communications purposes:

- Short-range device-to-device links between smartphones, tablets, TVs, PCs or IoT products, using their normal speakers and microphones. Data-over-sound competes with – or sometimes complements – WiFi, NFC, ZigBee, Bluetooth and even QR codes for "proximity networking". This is the primary focus of the applications described below.
- Acoustic watermarking and fingerprinting, which embeds extra data into music or speech audio. This can be used for purposes such as copyright-protection, or integrated classification of different content streams (eg musical genres). This is covered as a secondary domain here.
- Use of sound or (more often) ultrasound for presence-awareness, user positioning and movement sensing or gesture-recognition. This is essentially data-*from*-sound rather than data-*over*-sound. While this is a very interesting area in its own right, it is not the primary focus of this report.

Sound-based payments

A major focus area of innovation (and in-market deployments) for data-over-sound is for payments, especially from smartphones to retail payment terminals/kiosks, or from bank ATMs. The general theme is for secure, proximity-based links between customer and supplier, often linked to some form of "mobile wallet" on the device. Sound-based payments can be directly integrated into mobile apps, with nothing more than microphone/speaker access, and without dependency on the underlying OS or specific device.

While many developed markets have focused on NFC (near-field communications) chips, these have limitations: there are several proprietary NFC payment ecosystems; not all low-to-midrange phones have the capability; not all payment terminals support contactless; and the user-experience and interface can be awkward. Uptake is low and growing slowly, or even flat-lining. And in countries such as India or those in sub-Saharan Africa, many smaller merchants cannot afford card-reader terminals at all, let-alone ones capable of NFC or 2D-barcode scanning.

While international interest in sound-based-payments has been high, the concept seems to have gained most traction in India – a very mobile-centric market with low bank-account penetration, and a strong national ID system called Aadhaar.

Importantly for markets such as India, sound-based payments do not need always-on Internet connections in order to work. This is important in rural areas, or for users who pay per-day/hour for mobile data. Another advantage of DoS for payments is that it may not even need a smartphone – potentially it can be implemented in any device with a speaker/microphone, including wearables or even a plastic card. This reduces both the physical space requirements for NFC support, and the licensing / chipset costs.

As well as payments *from* mobile phones, DoS can also send virtual cash *to* the user's wallet from a bank ATM using sound or ultrasound. This is an alternative to either NFC-based "taps" on the machine, or cumbersome one-time passwords that must be entered on the device.

Numerous technology suppliers and bank/payment providers have entered the market for sound-based payments, including:

- Microsoft demoed a solution called Dhwani in 2013, but never launched it commercially.
- An early Silicon Valley innovator called Naratte, with an ultrasound technology called Zoosh, announced itself in 2011 and was subsequently acquired by payment specialist VeriFone. In 2014 VeriFone launched the technology for passengers paying for taxis in New York and Philadelphia, within an app called Way2Ride. (That taxi-hailing app has subsequently been merged with one called Curb, which doesn't support DoS at present)
- Alibaba's AliPay Wallet first trialled a mechanism to connect phones with vending machines (initially metro ticket-terminals) in 2013[2]. It is still listed on the website as a payment option.

[2] https://techcrunch.com/2013/04/14/alipay-launches-sound-wave-mobile-payments-system-in-beijing-subway/

- ToneTag (owned by Naffa Innovations) which works with Tech Mahindra's MoboMoney, Yes Bank and various others. It had signed up 25,000 merchants to its platform by the end of 2016.
- Other India-focused sound-payment startups include Paytm, iKaaz and UltraCash
- Soundpays is a Canadian startup using sound to enable proximity payments and also one-click online payments from TV broadcasts, outdoor display adverts or webpages with audio embedded into it.
- Amdocs, a large software company best-known for its billing software for telcos, has integrated data-over-sound into its mobile payment solution for banks and service providers. It has worked with a number of underlying technology providers for different types of proximity payment.

Industrial & IoT

There is currently huge growth in industrial automation of all sorts, ranging from distributed sensor networks in agriculture, to semi-autonomous robots and drones within factories. Ultrasound has long been used for sensing obstacles or movement. Unsurprisingly, there are parallel requirements for connectivity and data communications, both between IoT devices and controllers/servers, and in some cases between each other (for example to avoid collisions). Data-over-Sound has relevance for IoT/industrial use.

Obviously, there is a plethora of radio-based IoT networking solutions, at all levels of data throughput, range and power-consumption. 4G, WiFi, ZigBee, LoRaWAN, SigFox and numerous other wireless systems are available, with 5G on the approaching horizon as well. Yet in some cases, these are all still inappropriate – and sound-based networking may have an important role to play.

In particular, in some environments there are risks associated with RF – notably, where there are delicate electronic systems (eg in hospitals) or explosives (eg mining). In other situations, there may be sources of RF interference (eg welding machinery) or ionising radiation (eg in nuclear facilities).

In these circumstances, low data-rate, short-range transmission can be conducted with audible sound, or sometimes ultrasound where there is too much ambient mechanical noise. (Higher data rates may necessitate fibre connections, although these are obviously inflexible and expensive). Sound can also yield "intensity gradients", which robots can use for navigation, estimating distance from a source.

As with other uses of Data-over-Sound, the original idea of audio connections for robots has been around for some time – for example being described in a 2007 MSc thesis[3].

Consumer devices and electronics

Alongside industrial IoT, the other main area in which new electronic, interactive devices are being created is in the home, or other consumer environments such as in-car. Often, there is a need to connect new products such as TV set-top boxes, washing machines,

[33] https://pooyak.com/work/pubs/pooya-karimian-msc-thesis.pdf

smart utility-meters or doorbells, to secondary terminals such as smartphones or PCs, for example where the is a "companion app" or website involved for set-up and configuration.

It is particularly difficult to discover, connect and "pair" sets of random devices, ideally with "zero touch" on the part of the user. Long-lived devices like fridges or boilers may not share operating systems or even the same generation of wireless radio technology with smartphones, or even home WiFi controllers. One or more of the relevant devices may be provided by a service provider such as an electricity company or cable operator, or perhaps a healthcare specialist. Given the lifetimes involved for non-phone products, it is possible that users no longer have instruction manuals available. Moreover, one or more products participating may have limited or no screen/display (eg toys). Some customers may be uncomfortable or unfamiliar with IoT technology, and ill-suited to using Bluetooth or other options. Data-over-sound potentially improves the chances of successful device pairing, without friction or calls to customer support (or worse, a costly "truck roll" to the home for an engineer to visit).

An important category of consumer products, for which DoS has significant applications is in the area of toys and games. Increasingly, toys are adopting computing and "connected" capabilities, whether that is building blocks used for robotics, learning tools, racing cars controlled by mobile companion apps, or even soft toys which allow kids to speak to their parents remotely. A demo[4] of a Star Wars toy which "speaks R2D2" exemplifies the possible opportunities, where audible signalling is actually part of the end-user appeal and perceived value.

Another example is Google's Chromecast media player dongle, which connects to TVs via HDMI ports. A "guest mode" allows visiting Android tablet or smartphone users to control the Chromecast, without associating via Wi-Fi. The DoS system shares a PIN code to associate the two devices temporarily. So for example, a visiting relative could display cloud-resident vacation videos or pictures on a TV screen, without needing a cable, or cumbersome manual Bluetooth pairing or access to the WiFi password. A video demo is here[5].

We are also likely to see a lot of other audio-capable devices in homes in coming years, thanks to the trend towards voice-control and "chatbot" type products. While Amazon's Echo is the best-known, numerous other vendors and service-providers are also entering the market. US cable operator Comcast, for example, is intending to use its in-home devices and WiFi as a hub for smart-home services. At the moment, these products don't appear to use data-over-sound techniques, but given the sophisticated signal-processing used for the microphone arrays, this is a possible area for future evolution, possibly in combination with ultrasound-based proximity detection.

Retail and advertising

One of the most talked-about uses for DoS technology – and perhaps the most controversial – is using it for tracking consumers' online and offline activities. Typically using inaudible ultrasound signals, various companies have developed solutions to

[4] https://vimeo.com/channels/chirp/204443573
[5] https://www.youtube.com/watch?v=mHFf2M1Ce38

identify users entering – and moving around – retail stores. Other have focused on watermarking online ads or websites, and using smartphone apps to track users across multiple platforms.

Some use-cases have been legitimate, transparent and "respectable", for example loyalty apps collecting venue "check-ins" such as ShopKick (bought by SK Telecom for $200m in 2014). But others have been less-ethical. In particular, some (especially on Android) have had access to the microphone while the app was in the background, or even with a device in a user's pocket. This has threatened intrusive and privacy-invading location-tracking, potentially including "cross-device" tracking even where users try to keep their activities isolated. Further, even where apps have specifically been described as using ultrasound, the permissions for the app may allow it to listen (and analyse/record) *all* audio including local speech.

Sound / ultrasound is certainly not the only way retailers are watching and locating/identifying shoppers – they are also using Bluetooth beacons, face-recognition cameras, WiFi hotspots tracking device chipset "MAC addresses" and much more. Some such as ShopKick's use a combination of technologies. Nevertheless the use of APIs to create in-app ultrasound abilities has raised significant concerns over surveillance of consumers, especially in markets with strong privacy protections.

One of the companies involved, SilverPush, became particularly controversial in 2016, when it was discovered that watermarked TV adverts prompted some smartphone apps embedding its code to record personal data covertly, notably about viewing habits. The US Federal Trade Commission even issued a warning notice[6] about it. SilverPush subsequently pivoted away from its "audio beacons".

Several other vendors such as Mood Media and Signal360 have developed integrated solutions for retail/marketing tracking, while DoS generalists such as CopSonic and LISNR highlight this use-case in their marketing, although they vary widely in terms of actual in-market deployment. Other vendors such as Audible Magic, Shazam, Cifrasoft and Intrasonics can use audio watermarking to track users' exposure to advertising or other content for brands.

UCaaS / enterprise collaboration

Many audio-related innovations are driven by internal enterprise usage – especially by use in conference and meeting-room settings. Improving communications and collaboration effectiveness is something that businesses will pay for, whether that is for improving client-contact, or creating more effective teams. Video- and web-conferencing systems, digital whiteboards, and integration of mobile devices are all major trends, along with cloud-based "as-a-service" business models.

However, a perennial problem has been the usability of such "unified communications" (UC) systems. Most readers will have struggled to get a videoconference set up, or connect a laptop to share a presentation or file, for both the in-room projector and remote participants. Adding smartphones or tablets adds a new layer of confusion. A huge amount

---

[6] https://www.ftc.gov/news-events/press-releases/2016/03/ftc-issues-warning-letters-app-developers-using-silverpush-code

of time is often wasted – and costs incurred for tech support – by such "friction" during what is supposed to be a collaborative session.

Leading suppliers of UC and UCaaS solutions are now looking to create better ease-of-use, and ideally "zero touch" configuration, while maintaining good security/privacy.

Cisco is probably the most aggressive here, although PolyCom used acoustic pairing first. Cisco's Spark UC system (which is slowly replacing the older WebEx brand) uses an ultrasound pairing approach called Proximity. This can automatically wake up a white-board screen (the "Spark Board") when an app-enabled user enters the room and chooses a relevant action, allow sharing of documents or web URLs from laptops or mobile devices, or allow basic remote-control functions such as making and disconnecting calls. Because the user is already signed-in securely on their mobile device, the ultrasound pairing function allows frictionless second-screen access, without use of PINs or configuring Bluetooth (which is also supported).

A broadly-similar ultrasonic system from PolyCom called SmartPairing[7] has been available for several years, as part of its RealPresence portfolio. It claims various patents in this domain.

Conferencing and UC is also starting to embrace virtual or augmented reality. We may also find DoS being used to help pair or sync multiple users/devices in shared AR/VR sessions; this demo[8] by Visual Vocal is an early example.

WiFi access / geo-fencing

The WiFi industry is always trying to find easier and more effective mechanisms for authentication between users and access-points. While various forms of auto-logon have been developed, as well as portals allowing "manual" connection (plus monetisation from advertising), they all come with problems of user-experience, or access-control.

One of the issues with automatic WiFi logon is that signals can "leak" out of buildings onto the street outside, or through the wall to the next-door location. This can mean that some unauthorised people can connect (eg people sitting outside an airport lounge), or it may capture people just walking past the venue, rather than sitting inside.

Data-over-sound can help here – while WiFi signals can pass through walls easily, the same is generally not true for audio waves. A combination of normal automated logons *plus* the ability to detect a "sound SSID" inside the venue, could help control access. This is also applicable for other "geo-fencing" use-cases.

Public spaces & public safety

Various commentators and vendors have suggested that data-over-sound technology may be applicable for large-scale public venues such as sports stadia, or for use at

[7] http://www.polycom.com/collaboration/innovations/smartpairing.html
[8] https://vimeo.com/channels/chirp/207935997

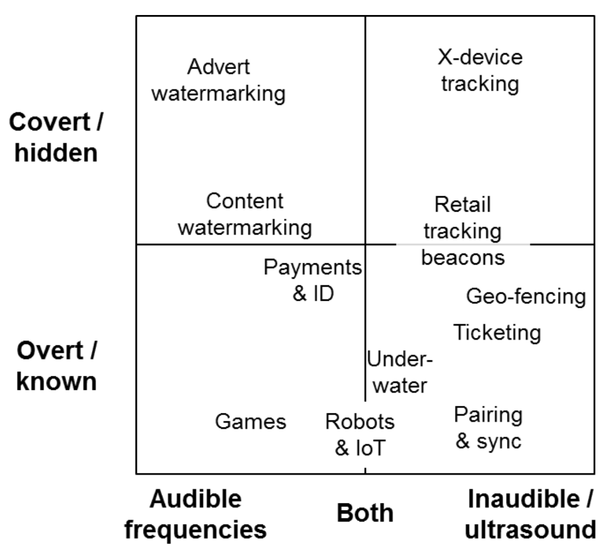
conferences/conventions, festivals or similar public events. US DoS vendor LISNR has claimed a number of trials and demos in this area, for example.

While there may be some workable applications, Disruptive Analysis remains somewhat sceptical. In general, DoS is not well-suited to outdoor acoustics over large areas, given issues around echoes (especially from concrete structures), ambient noise and the limitations of most PA systems. There are also often complex value-chains involving many stakeholders – site-owners, event-promoters, caterers, ticketing providers and so forth. Few consumers install event-specific mobile apps, so DoS may need to be integrated into whatever native-OS "wallet" is used. Very short-range acoustic signals may work for ticketing/admission, although it is unclear if this scales well to venues with many ticket-gates next to each other. For now, QR codes, which work on any device as well as printed paper, seem to be hard to displace.

Perhaps more interesting will be use-cases relating to public safety or national-security. It is possible that intelligence agencies may exploit some of the covert capabilities of ultrasound, for example.

Also, although it is not transmitting "through-the-air" data, it is worth noting that the new EU standard for in-vehicle emergency calls (eCall) uses ordinary 2G/3G phone calls, with an "in-band" signalling technique for sending data to public safety answering points (PSAPs). Basically, this harks back to the days of modems, by using a normal circuit—switched voice channel to send information – called MSD, the minimum set of data – such as GPS location, vehicle type & registration, fuel type, direction of travel and so on, to help emergency crews.

DoS use-cases span audible/inaudible, and known/unknown modes of operation



Source: Disruptive Analysis

Underwater communications

One of the more surprising domains where DoS has an existing strong role is actually in marine communications. Water absorbs most radio signals (except ultra-long waves around 30-300Hz), so there are only limited methods to send data, unless a physical tether cable is used.

For example, acoustic telemetry is used to send data from seabed pressure sensors to Deep-ocean Assessment and Reporting of Tsunamis (DART) buoys in the Indian Ocean, which in turn relay the information to a satellite. This forms part of a much broader tsunami-warning system, which also uses seismographs and other inputs. Other uses of underwater acoustic networking include communications with human divers, or remote-control of aquatic instruments, robots and drones.

Commercialisation pathways for Data-over-Sound

Although the previous section highlights some real applications of data-over-sound technology, it is important to acknowledge that its use is still far from mainstream today. Indeed, given that the basic concept of DoS is decades old, it is reasonable to question whether it can move beyond its historic niches, through tightly-customised applications, to become as much a generic and mainstream device-to-device communications tool as, say, Bluetooth or QR-codes.

In Disruptive Analysis' view, it is not enough for specialist systems-integrators to embed DoS into vertical solutions, though that has scope for very considerable short-to-medium term growth.

In the longer term, the technology has to be "democratised" to the extent that average app-developers or IoT device-manufacturers can adopt it, without huge investment in new skills. Eventually, it could become an inherent part of key operating systems such as iOS, Android or Windows – or be enabled directly in IoT chipsets and cloud platforms.

That end-goal, however, comes with caveats. Firstly, it needs to be recognised that acoustics are *hard* and dependent on the physical environment that a given application will work in. Secondly, DoS has a range of privacy and security implications that will need to be managed properly, by end-users and solution providers, at least until robust standards and codes-of-conduct emerge. Thirdly, awareness of, and comfort with DoS will need to percolate through consumers, business-users and perhaps regulators and governments. This will take time, and a number of intermediate steps.

In some ways, DoS is similar to another area that Disruptive Analysis has looked at in the past – the embedding of realtime voice and video communications into apps and websites, using technologies such as WebRTC. It is a very good underlying concept, with a variety of tools, APIs and PaaS approaches to enabling integration. However, often specific implementations still require extensive tuning and expertise of media, user-interface design and various other context-specific considerations. The lack of awareness of the benefits, and the lack of skilled engineers to implement it, has meant that uptake took longer than some expected – and concentrated it in major vertical applications rather than as easy-to-use horizontals for the "general" developer community.

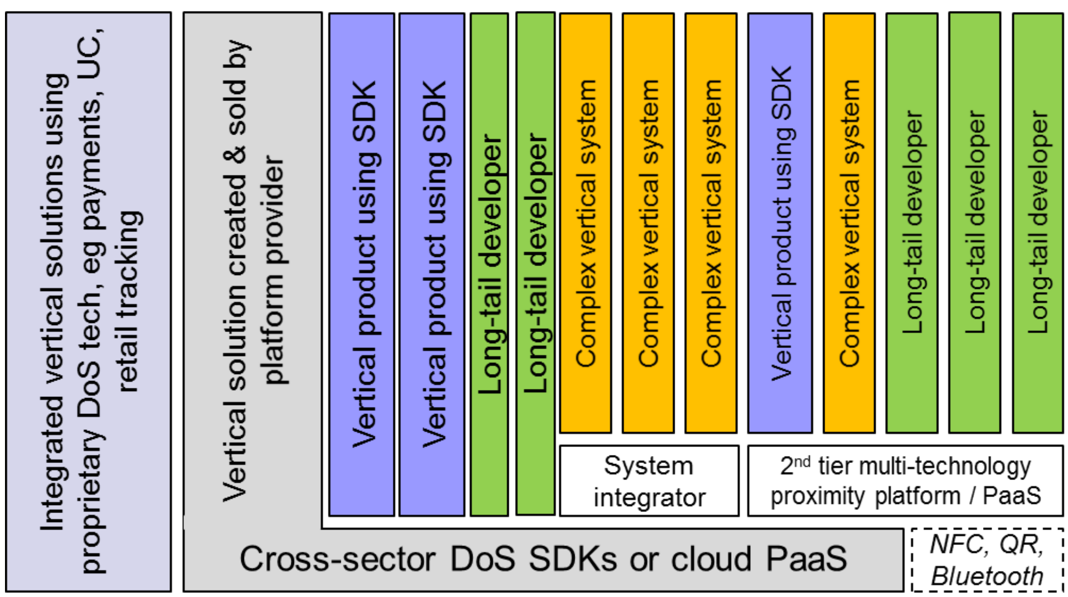
The rest of this section looks at what is needed for the DoS industry to short-circuit the normal market development process, and avoid potential delays and pitfalls.

Business models

There is no "one size fits all" business model around DoS – which is largely unsurprising, as it is a "capability" rather than a market, service or end-product in its own right. Its applicability to hugely diverse application areas and device types also tends to force it into multiple separate value-chains, at least for now.

The most obvious divide is between companies that are trying to focus on specific verticals - for example, sound-based payments, or in-store/in-advertising beacons - versus those aspiring to be horizontal enablers of many different use-cases. Then there are systems integrators which take those components, add extra layers for security, identity and so forth, and include them in broader products or solutions.

Data-over-Sound Value Chain: from standalone solutions to developer platforms



Source: Disruptive Analysis

There will also be intermediate-tier solution providers, which incorporate DoS into a broader platform, for example a PaaS play, offering multiple proximity-messaging tools to other developers. Unified Inbox is an example here - it uses Chirp's sound technology as one of many ways to provide human-to-machine messaging, which is then bundled into overall IoT offerings by device suppliers or larger-scale integrators or cloud operators.

Over time, this means that DoS is likely to manifest in 3 main ways:

- Integrated "behind the scenes" in point solutions
- Distinct and well-defined verticals (such as payments) which are explicitly sound-based. These will feature a mix of in-house development and 3rd-party elements
- Horizontal platforms and SDKs, feeding into long-tail and custom-developed solutions over a broad set of markets, especially for consumer and industrial IoT. In some cases there may be two or more layers of such platforms.

In terms of pricing, this structure means that DoS capabilities will be offered to customers (both B2B and indirectly as B2B2C) in various different models:

- Per-event (eg per sound-based transaction)
- Per-user / per-device on a subscription basis (per month/year)
- Per-endpoint / per-app one-off license
- IPR / patent licensing
- Systems integration or managed-services fees
- DoS-embedded services (monetised based on the relevant vertical overall)
- Ancillary services (storage, analytics etc)
- Non-charged (eg open-source)

Over time, we may see the emergence of major players that reshape the market around themselves, in the same way that Amazon has helped define cloud-computing and IaaS/PaaS models. However that is probably at least 2-3 years away for data-over-sound, and perhaps longer depending on interim developments and successful "cut-through" use-cases emerging.

Privacy and security considerations

While growth of DoS seems to be at the start of a possible exponential rise, it is also prudent to consider possible downsides. One of the risks for the DoS marketplace is around security – especially if it catalyses regulatory involvement.

The last 6 months have seen a sudden rise in the perceived privacy risks around sound-based communications, in particular around covert use of ultrasound. A number of academic studies[9] have identified how various technology and marketing companies have used ultrasound for "device fingerprinting" or "cross-device tracking". The main intention is to "de-anonymise" users to allow better targeting of advertising. This follows on from the 2016 Federal Trade Commission warning about covert tracking of user behaviour, especially around SilverPush's APIs.

In Disruptive Analysis' view, the key term here is *consent.*

Just as in social situations, or more prosaically for other device capabilities exposed via APIs, it is important for applications to be transparent and open with their users. It is critical that users understand not just bland statements ("Grant access from this app to the microphone"), but also understand *when* that capability will be used – for example only

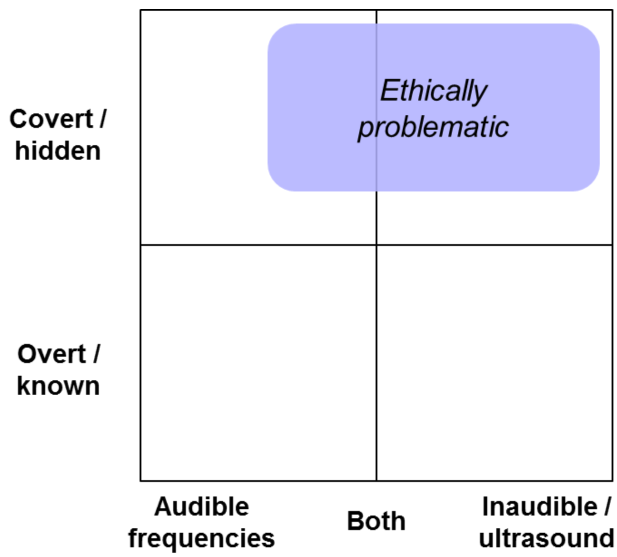
[9] https://www.cs.ucsb.edu/~vigna/publications/2017_PETS_Ultrasound.pdf

while an app is running in the foreground vs. background – and most importantly *why*. Ideally, the device operating system or app will give the user fine-grained visibility and control, such as only permitting access to ultrasound noises, and giving assurances that actual speech cannot be surveilled or recorded.

With audible sounds the user is at least always aware that their device is "talking" or listening to another. With watermarked or ultrasound communications, that is not always the case. A lack of "candour" that may limit the uses for ultrasound, especially when used by unknown / untrusted providers of services. Human psychology means we are very wary of communications being conducted "behind our backs" – an audible tone gives reassurance that we are not being spied upon. (This is similar to the reason many digital cameras have a shutter sound, or webcams have a green LED illuminated when active).

There are plenty of applications where ultrasound *is* trustworthy (for example, device-pairing for conference rooms, where it is a clearly-stated system feature and the apps can access microphones for normal voice as well). But some of the marketing/advertising applications are on shakier ethical ground.

Covert use of DoS technology poses issues for privacy & accountability



Source: Disruptive Analysis

The misuse of DoS risks additional regulatory scrutiny, wariness among the general public, countermeasures from OS suppliers and anti-virus/malware specialists, and a reluctance for PaaS models to be deployed fully. The industry needs to be seen to take this issue seriously – ideally with clear codes of conduct, or certification by a standards body.

Decoding vendor hype vs. reality

[Note: Disruptive Analysis does not do formal product evaluations / comparisons]

The vendor landscape for DoS is quite turbulent. There are large numbers of providers that have announced technologies, products or prototypes, but considerably fewer that have yielded commercially-proven deployments. The most notable market participants are those with specific point solutions adopted with big-name partners, or with their own go-to-market propositions.

Some well-known major companies have experimented with data-over-sound in their labs, but have yet to launch commercial products – for example Disney[10] and Microsoft[11]. Others such as Cisco and Google[12] have used sound/ultrasound for certain features around device-pairing or proximity-messaging, but are not monetising it directly.

Some smaller vendors appear to be adopting a "fake it until you make it" strategy – over-promising on their current capabilities, in the hope of attracting capital and attention in the short-term, backfilling in development afterwards. This is not unique to the DoS arena – Disruptive Analysis has seen similar hyped claims in many other fields of innovation. Customers, investors and commentators should look to real-world demos and deployments, and maintain a wary scepticism of "outliers". Some warning-signs include claims of:

- Very long-range DoS applications. The physics of sound propagation in air mean that distance is a huge constraint. Audible and ultrasound signals are subject to an inverse-square power law; are affected by wind, rain and ambient noise; and reflected, scattered and absorbed by the environment. Distances cited over 10-20m are likely to be problematic, except perhaps over loud venue PA systems.
- High-speed data transmission. Claims of kilobit+ data-transmission speeds for DoS are highly dubious, given the width of frequency ranges involved, acoustic distortions, echoes and speed of sound in air, the need for repetition of signals for error correction, and the computational requirements for signal processing. At very close range (mm or cm) or with specialist equipment, higher speeds are feasible, but for general room-scale use any such claims need careful testing and validation.
- Outright replacement of Bluetooth, NFC or QR-codes. In general, DoS is a complement to other short-range technologies, not be a direct substitute. It is extremely unlikely that DoS will render any of those obsolete. Most obviously, Bluetooth (& BLE) is much larger in adoption-scale and data-rate, while QR codes are even more versatile than sound as they can be implemented on paper as well as screens. NFC is perhaps more threatened, but even that is linked to popular contactless-card technology and is unlikely to disappear. For some use-cases, DoS will be better – or used in combination with one or more of these. But those citing it as a revolutionary replacement are over-selling it.
- Ability to support multiple simultaneous DoS users in a confined area. Even leaving aside the aesthetic impact of continual, overlapping audible "bleeping", there are

[10] https://www.disneyresearch.com/project/mobile-phone-arrays/
[11] https://www.microsoft.com/en-us/research/publication/dhwani-secure-peer-to-peer-acoustic-nfc/
[12] https://developers.google.com/nearby/messages/overview

likely to be significant acoustic interference effects – which also apply to (near) ultrasound as well.

- Commercial adoption without clear evidence: Actual deployments typically have independent validation such as 3rd-party product reviews, statistical data on user base or engagement, unambiguous press releases or in-field videos and so forth. Proofs-of-concept, lab demos and PowerPoint slides are all interesting, but need to be viewed with caution and not extrapolation.
- References to TV broadcasts using DoS are questionable, especially for inaudible/ultrasound messages. While (some) TV speakers may be capable of DoS, there may well be various forms of filtering and compression used in the network which strip it out, or render it error-prone. Most network media-server and transcoder boxes are not designed to deliver or protect ultrasound.

It is also worth noting that the most "press-friendly" vendors are those with consumer-facing propositions around payments, retail or consumer IoT. There is less incentive for systems integrators working on industrial robots to look for media coverage, although their component or platform providers may seek to use them as case-studies.

Lastly, it is worth being wary of "ghost" vendors. Some companies have made few announcements about customer wins or product evolutions in some time (years, in some cases) – and while this doesn't mean they have stalled or disappeared permanently, it should certainly spark some questions if developers are considering them as mission-critical suppliers or platform providers. Even Google Nearby Messages has very little documentation or discussion about the ultrasound API, compared with its Bluetooth peer[13].

Overall – while DoS is "real", not all the claims around it are, so: "*Caveat Developer".*

Building the ecosystem

If data-over-sound becomes established as a key part of the Internet and IoT landscape, we can expect numerous extra opportunities to emerge as it matures. Indeed, most of these will be *necessary* for DoS to become a mainstream, "everyday" part of the app and IoT landscapes – without them, the technology will be restricted to a smaller set of custom-engineered vertical solutions.

- **Test and measurement:** Any emergent technology requires effective test mechanisms in order to reach mass-market adoption. This includes both lab/pre-deployment assessment tools, and methods for assuring ongoing performance. DoS will need its own test products, especially if it is going to be used by mainstream developers and designers, who have limited experience of acoustics.
- **Training and skills**: At the moment, only a small number of developers and engineers are familiar with DoS as a concept. Few software or IoT designers have acoustic skills. While it is unrealistic to expect every organisation to hire PhD-level specialists, there will need to be at least rudimentary training resources. Events such as hackathons may be useful, as well as online courses and support forums.

[13] https://developers.google.com/android/reference/com/google/android/gms/nearby/messages/audio/AudioBytes

- **Standardisation**: While DoS implementations will continue to need context-specific tuning, it is likely that some underlying aspects will need official standards. In particular, web browsers and maybe key operating systems will be wary of supporting proprietary acoustic-communications protocols. Some form of DoS industry forum or trade association would also help, in terms of driving awareness, defining ethical policies, and also maintaining interactions/requirements with regulators, standards bodies and related organisations.
- **Platformisation:** At the moment, DoS implementations are mostly "artisanal", requiring skill and expertise either by a specific vendor, or a small number of specialised integrators. While APIs and SDKs for apps are available, they still need careful optimisation. For the technology to spread more widely, it will need to be embedded more accessibly into popular platforms – which could be device OS's, cloud-based PaaS, IoT frameworks – or even directly into chipsets.
- **Open-Source:** In order to develop DoS trials and proofs-of-concept, many developers will need a "kick-start" to encourage experimentation. In many other areas of technology, this now involves easily-accessed code libraries, and open-source building blocks. This does not preclude the use of paid software or platforms, especially as use-cases scale up. But reducing friction for applications to get off the ground is very important.
- **Data and analytics:** The key source of value in many areas of business is going beyond transactions, towards analytics and machine-learning. Capturing data about the use of DoS, its effectiveness in helping users or devices, and changing behaviour or application outcomes, will increase in important. There may also be regulatory and compliance requirements in some sectors to store audio DoS messages.
- **Hybrid platform management:** Various use-cases will combine DoS with other device-to-device technologies such as Bluetooth or QR-codes, to give as broad as possible coverage of user preferences, device support and application context. There will likely be subsystems or integrated software platforms to support "multi-channel proximity communications".
- **Interference management and mitigation**: As DoS becomes more widespread, the probability of two or more applications running concurrently on the same device, or in the same physical space, will increase. There is a likelihood that this will cause problems with acoustic interference, or software conflicts with different apps trying to access microphones and speakers simultaneously.

We can still expect DoS to expand in use without all of these extra ecosystem dimensions. These areas will not appear overnight, in any case. Yet for DoS technology to become a resounding success and become a "democratised" method for device communications, paralleling Bluetooth, QR-codes and others, we will need a lot more "enablers" beyond the basic solution components.

Roles of telcos, SPs & cloud players

At the moment, telcos have a very limited role in the data-over-sound landscape. Retail solution ShopKick, which includes an ultrasound component, is owned by South Korea's SKTelecom, although more because of its synergy with online/offline commerce than its voice/audio elements. In theory, there should be interest in the domain, especially as the industry pushes higher-fidelity codecs into mobile and fixed networks for HD voice calling. However in reality, most "mainstream" units of telecom companies do little beyond traditional voice interactions such as telephony and conferencing.

Perhaps more relevant are the units of service providers (SPs) that focus on IoT enablement, smart-home services or mobile payment solutions. Often, these are peripheral from the "traditional" network-centric departments of the telco, and more able to consider on-device capabilities that may work "offline".

It seems more likely that DoS will attract attention from other groups of service providers that are less telco-centric. Google, Amazon and Cisco are the obvious participants here, with (for example) Cisco already integrating ultrasound into the proximity function of its cloud-based Spark unified communications platform. Google uses ultrasound in its Nearby proximity API, although its developer literature seems to downplay this in favour of Bluetooth and other capabilities.

Amazon is also aggressively turning its Echo/Alexa voice assistant into a broader platform, linking in-home equipment to cloud-based capabilities and (perhaps) APIs. It is likely to continue evolving its Echo and Dot, which already include up to 7 microphones and considerable acoustic know-how. Recently, Amazon has made a microphone kit available to external hardware developers. If it enhances its cloud-based software platform to support DoS in future, it should have a large and growing installed base already active.

At the moment, there is no real equivalent of an Amazon AWS or Twilio offering full Internet-based PaaS options for data-over-sound. It is not possible simply for a developer to just sign up with a credit card for a cloud-based DoS (DoSaaS?) analytics/signalling capability, or embed generic APIs into a mobile app for device-centric functions. The heavy amounts of customisation required, and the background reputational risks of privacy-invasive applications, make it hard to have broad horizontals at this stage of market development.

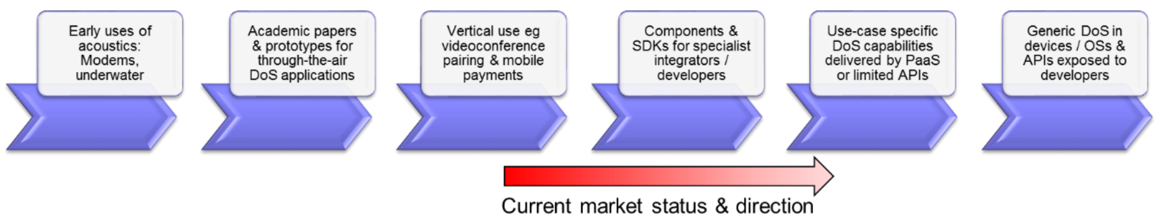
Disruptive Analysis expects narrower PaaS offers and open SDKs to emerge for particular verticals. It seems quite likely that industrial IoT will aim to standardise certain aspects of DoS, or else perhaps strong IoT platforms or hubs may offer certain functions for a limited audience. Secure payments using sound could also be "platformised" by cloud players.

Conclusion & recommendations

Data-over-Sound is not a new concept. Yet until recently, it has had very limited adoption beyond narrow niches. The proliferation of smartphones and IoT devices is changing this rapidly, with important and fast-growing opportunities in consumer IoT (such as toys and "on-boarding" new appliances to smart-home systems) and industrial automation. Mobile payments and retail/marketing use-cases are also important niches.

However, it is still easy to over-estimate its level of maturity. It seems hugely valuable for applications such as mobile payments and device-pairing, or usage in specific environmental contexts – but at the same time, it needs heavy customisation and the underlying skillsets of acoustic engineering are rare commodities.

Data-over-Sound Market Development & Trajectory



Source: Disruptive Analysis

It is worth remembering how slowly and patchily adoption of technologies such as NFC and QR-codes have been for "proximity" communications. Bluetooth really only gained a rapid foothold because of a widespread need for phone headsets and in-car connections – as well as having an effective industry body for standardisation and promotion.

This means that existing market participants need to start working together now, to catalyse the overall health of the sector for its common good. It would be beneficial to see something like a "Data over Sound Forum" emerge, which could establish working-groups on standards, ethics and perhaps certification – as well as publicity and as interface/advocacy point for governments and regulators. Vendors should look to build links between each other, even if they are bitter competitors. This is a typical technology niche in which a rising tide should float all boats – even if some are slightly less sea-worthy than others. The work of the Bluetooth SIG (special interest group) is perhaps a good model to emulate.

One of the challenges in assessing data-over-sound opportunities is its position at the intersection of many separate disciplines: smartphones, IoT, acoustics, identity/security, telecoms and user-experience design, to name just a few. It is likely that the next few years will yield significant positive surprises, and also setbacks.

In many ways, DoS exemplifies a wider renaissance of audio and acoustic technology. There is huge innovation not just in voice (for example, around voice-control and analytics) but also audio-processing from arrays of microphones, or the use of ultrasound. As well as data transmission, we should also expect a greater array of motion and location-sensing, gesture recognition, surface/texture measurement and even energy-harvesting from microphones.

As a recommendation, Disruptive Analysis encourages developers and device suppliers to experiment with data-over-sound, create prototypes and proofs-of-concept, and look perhaps towards audible tones as a way to "humanise" device interconnection, as well as exploiting the background invisibility of ultrasound. While we wouldn't want a cacophonous future of 100s of items bleeping around us, there is a positive reassurance sometimes that something is "working" – hence the common use of keypad tones and other audio design elements.

On the vendor side, companies should consider paths to the long-term future of the sector, and how it might be democratised across a wider developer-base, without cannibalising realistic revenue streams. A major obstacle is likely to be a shortage of skilled acoustic specialists, so both training and easy-to-use tools will be needed. Look to evangelise audio connectivity through hackathons, developer competitions and broader "outreach" at all manner of mobile-app and industry-vertical events.

Do not view the emergence of open-source elements as a threat – it is useful as a way to encourage experimentation and trials, with companies likely to move to commercial offers as they scale up.

Lastly, it is important for the industry to "create a category" for itself, either with the term "Data over Sound" or something similar. This is important to drive awareness among developers, investors, integrators – and ultimately end-users as well. There is big difference between technologies which are just component "enablers" versus ones that can garner a broader profile and momentum – for example, AI or speech.

About Chirp

Chirp is a fast growing B2B technology company. Its data-over-sound technology is founded on years of unparalleled scientific research and continues to be developed by some of the world's leading acoustic technologists and engineers.

Delivering both audible and inaudible options, its suite of interoperable SDKs work on many different technology platforms including iOS, Mac OS, Android, Windows UWP, JS, Web Browsers, Python, Arduino, Raspberry Pi and Linux. It develops custom SDKs for partners' specific requirements where needed. With Chirp, companies now have a technology that enables the interconnection of millions of new types of things in a seamless, scalable, cost-effective and effective way.

.

chirp.io

About Disruptive Analysis

Disruptive Analysis is a technology-focused advisory firm focused on the mobile and wireless industry. Founded by experienced analyst & futurist Dean Bubley, it provides critical commentary and consulting support to telecoms/IT vendors, operators, regulators, users, investors and intermediaries. Disruptive Analysis focuses on communications and information technology industry trends, particularly in areas with complex value chains, rapid technical/market evolution, or labyrinthine business relationships. Currently, the company is focusing on 5G, NFV, IoT networks, spectrum policy, operator business models, the Future of Voice, AI, blockchain & Internet/operator ecosystems and the role of governments in next-generation networks.

Disruptive Analysis attempts to predict - and validate - the future direction and profit potential of technology markets - based on consideration of many more "angles" than is typical among industry analysts. It takes into account new products and technologies, changing distribution channels, customer trends, investor sentiment and macroeconomic status. Where appropriate, it takes a contrarian stance rather than support consensus or industry momentum. Disruptive Analysis' motto is *"Don't Assume".*

For more detail on Disruptive Analysis publications and consulting / advisory services, please contact information@disruptive-analysis.com. For details about Future of Voice & Contextual Communications workshops & publications, please see www.deanbubley.com.

Website: www.disruptive-analysis.com **Blog:** disruptivewireless.blogspot.com
Twitter: @disruptivedean **Quora:** Dean-Bubley

Intellectual Property Rights / Disclaimer

Every reasonable effort has been made to verify research undertaken during the work on this document. Findings, conclusions and recommendations are based on information gathered in good faith from both primary and secondary sources, whose accuracy it is not always possible to guarantee. Disruptive Analysis Ltd. disclaims all warranties as to the accuracy, completeness or adequacy of such information. As such no liability whatever can be accepted for actions taken based on any information that may subsequently prove to be incorrect. The opinions expressed here are subject to change without notice.