November 2018
Geoff Huston

Internet Economics

One year ago, in late 2017, much of the policy debate in the telecommunications sector was raised to a fever pitch over the vexed on-again off-again question of Net Neutrality in the United States. It seemed as it the process of determination of national communications policy had become a spectator sport, replete with commentators who lauded our champions and demonized their opponents. We've moved on from that clamour, albeit without a very satisfactory resolution, and the dominant issue at present is all about the roles of the unabashed titans in our midst (Facebook, Amazon, Apple, Microsoft and Alphabet, assuming you have been living in a cave for the past decade and you need a quick list to refresh your memory!). Are they now so big they are essentially answerable to no nation state at all, or can we create a regulatory framework that places the interests of these technology behemoths into a more even balance with various national public policy objectives? This is perhaps the biggest public telecommunications policy question of late 2018, but by no means the only one. The other public policy conversations are not going away, including the topics of security and personal privacy, peering and interconnection, market efficiency and continued innovation and evolution in the communications domain, data-driven policy making, public measurements, and consumer protections to name just a few.

The way in which we communicate, and the manner, richness and reach of our communications has a profound impact on the shape and function of our economy and our society, so its perhaps entirely proper that considerations of the manner in which we develop and tune public policies in this industry take place in open forums. How can we assist and inform that consideration?

One way is to bring together the various facets of how we build, operate and use the Internet and look at these activities from a perspective of economics and public policy. This is the background to a relatively unique gathering, hosted each year by CAIDA, the Centre for Applied Internet Data Analysis, at the University of California, San Diego, at WIE, the Workshop on Internet Economics. These are my thoughts from the 9th such workshop, held in December 2018. The workshop assembled a diverse collection of individuals, including network operators, service operators, regulators, economists, legal experts, researchers and academics and the discussions were insightful and, for me, informative and helpful.

In setting the scene for this workshop, we are seeing the rise of a more sectarian and segmented environment, coupled with the re-imposition of various forms of trade barriers that appear to work in a manner contrary to the ideals espoused in a world of accessible open communications that is open to all forms of competition and innovation. The traditional view of a public space in the communications realm forming part of a society's inventory of public goods is being challenged by rapacious private sector actors who are working assiduously to privatise this formerly public space in ways that are insidious, sometimes subtle and invariably perverted. Even the surveillance of our individual profiles and preferences and the assembling of billions of personal dossiers is now being achieved by the private sector in ways that would make many police states from former days highly envious. What should be the role of government in this environment, and what can and perhaps should be assigned to private sector market actors to determine is still very unclear.

From an historical perspective, the disruptive changes that this rapid deployment of information technologies are bringing to our society are very similar in terms of the breadth and depth of impact to those that occurred during the industrial revolution, and perhaps the thoughts of the nineteenth century economist and revolutionary sociologist Karl Marx are more relevant today than ever before. We are seeing a time of uncertainty when we are asking some very basic questions: what is the purpose of regulation and whom should

we apply it to? How can we tell the difference between facts and suppositions? What are the tenets of the world that is being shaped by this digital revolution?

The initial euphoria of the boundless opportunities created by the Internet are now being balanced by common concerns and fears. Is the infrastructure of reliable and trustworthy? Why are we simply unable to create a secure Internet? Why do we continue to pass more roles to automated systems while at the same time we recognise that we are increasing our vulnerability to systemic failure modes that we are incapable of defending against? What is really shaping the Internet's development and is that evolution of the Internet's market forces creating a narrow funnel of self-interest that is driving the broad agenda? It seems that the Internet is heading down the same path as newspapers, where advertising became the tail that wagged the dog. In the case of the Internet it seems that we are being driven by surveillance capitalism, where intimate knowledge of each and every consumer, their desires and preferences and most importantly their spending habits are of obsessive interest. What is "private" in such a world? And do we all see the same Internet? For example, Facebook undertakes a deliberate effort to customise the platform to match as exactly as possible the preferences and desires of each user. Each user sees a customised social environment is the deliberate and intended result. Is this happening in other parts of the network as well? Are we now only seeing what these systems believe is what we each want to see? As these network content operators hoover up more and more of our profile data do we as consumers have any realistic valuation on the personal information we are exchanging for access to digital services? Is this a fair exchange of value? What's the ultimate social cost of a free search service?

With all that in mind, the context is set for this year's workshop.

Security, or the lack thereof

The aviation industry has much to teach us about how to manage security. If we fail to learn from our errors in operating complex machines, then we are doomed to repeat the errors. We need to know clearly when events happen, what happened and why it happened. Only then can we learn from these events and improve the safety and security of our environment.

There are elements of mandatory reporting in some areas and in some regimes but it's by no means uniform and the issues about potential liabilities incurred by service operators tend to encourage a certain level of reticence in mandatory disclosures. Harsh penalties, such as those used in the European GDPR framework have the unintended side effect of increasing the barriers to market entry for smaller sized players, for whom any breach would result in a fine of unsustainable proportions. The punitive nature that appears to lie behind the imposition of such significant penalties does not encourage an environment of full and open disclosure of all security events irrespective of whether a formal breach has occurred, so the opportunity to analyse, understand, and improve the quality of our digital infrastructure is forgone. I was struck by the observation that one's privacy of personal medical history is respected most of the time, but if you report to a hospital with a highly contagious and dangerous disease then a new regime of mandatory reporting swings into action. The imperatives of protecting all of society from such potentially devastating outbreaks of disease outweigh the normal conventions of individual privacy. Why isn't there a similar code of disclosure for data breaches? What's the appropriate policy settings that would allow us to understand the threat landscape and take appropriate measures to defend our data?

And when considering the defensive capabilities of our digital infrastructure, it's interesting to observe that these days we "stress test" our major financial institutions. We understand the critical nature of their function so much so that "too big to fail" is taken to the point of stressing the institution to understand if there is a failure point and at what level that may occur. But do we treat our data handlers in the same was as our finance handlers? Are we willing to conduct public stress tests on data integrity and resilience in the part of these providers? The comment was made that: "We are quite capable of building systems that we are incompetent to operate."

Underneath this is the perspective of a data-driven policy process, and the question is what measurement data would help us understand the resilience of our digital infrastructure? Easy measurements are not necessarily informative measurements, while informative measurements are often complex, challenging and costly. Perhaps

one critical question is whether the environment of security breaches is getting better or worse. How do we compare the breach of 500 million user credentials from a hotel chain's loyalty database against the breach of a single individual's credit card? Should we try and create some form of impact metric of security breaches that attempts to weigh both the severity of the event and the number of impacted individuals?

And how do we defend ourselves, irrespective of whether the primary cause of the failure or breach is the result of some failure mode within the system or the result of some form of deliberate attack on the system? Part of the reason why digital defensive capability is such a difficult problem lies in the complexity of many of our digital systems. In such systems it could well be the case that we can make individual components secure, but does that result in a secure system? Such systems often have complex interactions and these interactions can create brittle system behaviours. Without a clear understanding of the vulnerabilities, system defenders can't decide where to focus attention. What is the most important asset to defend given finite defence resources? And how can we tell the difference between adequate defensive postures and over-achieving? How do we determine when we are at the point where further effort would yield only marginal improvement in defensive capability?

And how do we orchestrate these defensive capabilities? Is this as public function funded by common taxation, or a private function funded by those seeking defensive capabilities. There is a world of difference between a conventional public policing function and a private security function. The public policing function is necessarily focused on apprehending the perpetrator and securing the environment to attempt to ensure that the event does not recur. The private security function is funded by the potential victim and the focus is to ensure that the potential to be the victim is realized. One could characterise the public function as "to try and make sure it does not happen again" and the private function as "to try and make sure it does not happen again to me." There is a world of difference between these two perspectives.

Experience with Security Practices

Security spans a broad range from the systemic overview to individual practices, and the workshop heard of the experiences with the introduction of two-factor authentication (2FA) in a community of users. Two-factor authentication has achieved a certain level of momentum these days and for many service providers there is the bandwagon effect where "if everyone else doing it, we should too!"

There is a question as to the effectiveness of this measure. To what extent does 2FA counter breaches? What is the breach reduction rate when 2FA is introduced into a system? By adding additional steps into the access process there is likely to be some level of fall off from consumers attempting to access the access the service? What is the drop off in usage?

The reported experience indicated that with a critical system, such as is associated with a personal and payroll system, they were unaware of any data breach since the introduction of 2FA into the system. For their email system the experience was a little different, in that they noted a visible drop off in usage with 2FA. Where the service is discretionary, the introduction of 2FA will act as a disincentive to some users. All of these systems represent some level of compromise between ease of use and potential to be compromised. The more challenging a system is to use, then the user response is to either drop off the system or where possible use short cuts that undermine the intent of the system challenges.

There is another class of security practice that is more indirect, and that is in the area of good network housekeeping with routing and addressing. One component of our vulnerability to attack is the distributed denial of service attack (DDOS) that uses co-opted systems to emit streams of UDP packets with the source address of the intended victim. There are a number of UDP protocols where the response is significantly larger than the query, including certain DNS queries and memcache requests. What this implies is that a small stream of trigger packets directed to conventional (and uncompromised) servers may result in a very much larger stream of data directed to the victim, potentially overwhelming their network. The key aspect of this particular form of DDOS attack is the ability to generate IP packets with a forged source address and have them passed across the network. We have largely given up on trying to prevent host IP stacks from generating such synthetic

IP packets, so attention has turned to the network. Can the network itself discard packets that have a forged source address?

The original concept was published as BCP38 (RFC 2827) some 18 years ago, and while the concept has not changed at all since then, network operators are still not enthusiastic to take the necessary steps to implement this. The issue here is that piecemeal adoption of the measures described in BCP38 are in themselves ineffectual in countering forged source address UDP-based DDOS attacks. For this measure to be effective, then all networks need to take these steps. CAIDA has been operating an active probing experiment to identify those networks that permit spoofed IP packets (http://spoofer.caida.org/). The results are somewhat mixed. The level of adoption of anti-spoofing in networks has not changed significantly for some years, while the observed incidence of UDP attacks using source address spoofing appears to continue unabated and has even increased over the same period. The use of tools that provide public disclosure of networks, using a so-called "name and shame" approach, has had some level of positive impact, but as we have learned with BGP de-aggregation, naming and shaming only has a short-term impact and the attention span to such lists is very short lived.

This could be characterised as a form of market failure where many individual network operators do not perceive sufficient self-interest to undertake actions that would prevent source address spoofing, and while this situation remains the case then the common good, namely a network that will not permit source address spoofed packets, is compromised. The result is that DDOS attacks using spoofed source addresses continue largely unabated. It is challenging to see a clear path forward here. If network operators incurred some liability for allowing otherwise preventable attacks from taking place, and if insurers were clearer about the impositions they would impose on operators before indemnifying them about such risks then perhaps the situation would improve. But this is an unlikely scenario in the near term. There is also the consideration of the adaptability of the attacker. We understand that attackers are opportunistic rather than rigid in their behaviours. Source address spoofing might be a convenient means of launching a DDOS attack today, but this approach is by no means the only attack vector and securing one attack vulnerability simply steers attackers to use other vectors.

This concept of the inability to secure a common system because of insufficient local incentives also applies to the picture of routing security. Sometimes long-term problems remain as problems because there is insufficient motivation to act to correct them. Sometimes these long-term problems are hard problems to solve because they require the coordinated actions of many parties, and the scale of orchestration required to field an effective response is simply forbidding. Sometimes they remain as unsolved problems because they resist our current understandings of how to solve them. Securing the inter-domain routing system appears to be an unsolved problem that sites in one of the latter two categories.

Reliable data to describe this issue is elusive. It's hard to distinguish between transient routing states, accidental misconfiguration of the routing system and the results of deliberate efforts to subvert routing. It's even more challenging to ascribe a degree of impact to a routing incident. A routing attack affecting a popular online service, or an attack that impacted on vital services that we all rely upon may well be considered to be a far more critical insecurity event than an incidental episode that might impact only my home service but nothing else. The issue here is that the routing system sees only address prefixes and paths and does not distinguish between them in terms of relative criticality or importance.

In observing that this is a long-term problem with a considerable degree of difficulty does not imply that no progress has been made. We have made progress in a number of areas. One part of the issue was the lack of a clear and testable authority model that allowed an IP address holder to attest that they are indeed the actual current holder of that address, and equally to be able to disprove all other contemporaneous claims of tenure of that IP address. We've devised a public key infrastructure (PKI) that allows IP address holders to digitally sign attestations about addresses and their permitted use and allows all others to validate such attestations. We have developed a standard technology model of how to incorporate these digital credentials into the operation of the BGP protocol and use these digital credentials in the operation of the routing system.

However, despite these advances, and despite the quite extraordinary effort to get to this state, we appear to hold little hope that this secure routing protocol, BGPSEC, will ever see universal deployment. There is a common view that the tool is too complex to deploy, and it use introduces a new set of operational risks that

outweigh individual network operator's perceptions of the risks of being impacted by routing attacks. This is coupled with the observation that the measure only addresses part of the problem space, so this measure would not secure the network from all forms of routing attack. It appears that it's a case where the cure might pose a greater set of risks than the original disease, and we are not even sure how good a cure it is in the first place!

Finally, there is the issue of the limitations of positive attestations in a secured framework. Positive attestations rely cooperating actors to be able to mark their actions or digital artefacts as *good* in some manner that providers clear attribution and inability to repudiate. This works if one assumes that a bad actor wishes to avoid clear attribution or is not in a position to obtain the credentials that generates these markings. In an environment where all the good actors undertook such marking at all times then unmarked material is clearly *bad*. But such environments of universal adoption are uncommon in massive distributed systems. In environments where the use of such marking of *good* is only performed by some actors, then it is unclear how to consider unmarked material. It is not necessarily *bad* or *good*. It is also not clear that such *good* marking capability can be withheld from bad actors. Phishing web sites use TLS and often display the green lock icon in the browser in the same manner as genuine web sites. Visible attribution and public reporting are often unhelpful as well. For example, Certificate Transparency is hopelessly ineffectual to counter short-lived "smash and grab" web attacks.

The general observation in this realm of secured systems is that imposing costs on *good* often does not eliminate the potential of *bad*, and if the total of these costs of marking *good* exceeds the loss of value through consequent theft and fraud than at a systemic level we have managed to introduce a net inefficiency into the system. In such circumstances it would probably be more efficient and cheaper to create a common loss compensation fund and carry the insecurity risk.

How then should we look at routing insecurity? Is this a failure of technology? If we had a better routing widget that could operate efficiently and cheaply then would we still have a problem? Or is this an instance of failure in the market? Individual network operators see insufficient marginal advantage in deploying these tools and the common advantage of a secure overall system is never realized. Or is this an instance of regulatory failure? If we have to turn to some form of regulatory imposition on network providers to secure the routing system, then how will we do this? The current national and regional regulatory examples in content control and encryption are woeful examples! Industry codes of practice have little effect if compliance is never checked and when consumers have no perception of the value of the code in any case. Why would comparable efforts in routing security fare any better?

The Changing Internet Economy

Annual global revenue within telecommunications services sector is variously estimated to be around 1.5 Trillion USD, which is slightly more than the global energy market and double the revenues in the aviation industry. While this represents considerable value in the telecommunications sector it is not disproportionately larger than many other major activity sectors, as the telecommunications sector represents just some 2% of the global GDP.

However, within the telecommunications activity sector the cost components have changed dramatically over the past couple of decades. The old Bell telephone service formula was a roughly equal division of service costs between the access network, switching equipment and long-distance carriage, and this general cost attribution model applied to all operators of voice switched circuit systems. These days the switching and long-distance carriage costs have dropped dramatically to essentially negligible levels. Packet switching is massively cheaper than time division multiplexing and circuit switching costs and fibre-optic carriage systems offer massive improvements in both the capex and opex of carriage systems. This massive reduction in switching and carriage costs have created natural incentives for the larger content and cloud providers to build out their networks all the way to the access networks. Compared to their other input costs it is inexpensive to do so and in so doing they remove a layer of intermediation between their service and their customers. This shift also deprives volume from the carriage provider sector, further fuelling their relative decline within this sector.

One could argue that the moves by some of these carriage providers to purchase content service providers may serve only to delay the inevitable terminal fate for these carriage providers rather than forging a completely

different outcome. As we have seem in many other enterprises when faced which transformational changes, such as the transformation of the transportation of goods from shipping to rail for example, the old generation of incumbents often do not have the organisational capability, the capital profile, adequate investor support and preparedness to openly concede that the risks to their core business are so inevitable and so terminal that would allow them to completely divest themselves of the past and reinvent the enterprise in an entirely new guise. Legacy considerations often generate perverse outcomes in times of business transformation. Carriage enterprises who invest in content service platforms often simply remain at their heart carriage enterprises that are now hosting a totally alien business unit within their enterprise.

Business transformation is often challenging. For example, a large multinational corporation started in Australia as a mining enterprise owned and managed by individuals who were essentially farmers. By the time they changed their internal management and investor profile they had become a steel enterprise owned and operated by miners. Another generation of change saw the corporation become an oil enterprise owned and operated by steelers. Then the oilers found themselves running an energy enterprise. At every stage of their corporate evolution, a major source of inefficiency and poor decision making was due to this misalignment of activity to the company's management and investor profile. It is tempting to see the attempt by these carriage enterprises to diversify in the same light as this example.

What we have today with the rise of content and cloud providers into dominant positions in this industry is a more complex environment that is largely opaque to external observers. What matters for consumers is their service experience, and that depends increasingly on what happens inside these content distribution clouds. As these content data network (CDN) operators terminate their private distribution networks closer to the customer edge, the role of the traditional service providers, which used to provide the connection between services and customers, is shrinking. But as their role shrinks then we also need to bear in mind that these carriage networks were the historical focal point of monitoring, measurement and regulation. As their role shrinks so does our visibility into this digital service environment.

It is a significant challenge to understand this content economy. What services are being used, what connections are being facilitates and what profile of content traffic are they generating, and just how valuable is it?

This brings into the forefront another venerable economic topic: is *big* necessarily *bad*? There is little doubt that the digital environment is dominated by a small number of very big enterprises. The list of the largest public companies as determined by market capitalisation includes the US enterprises Alphabet, Amazon, Facebook, Microsoft and Facebook and the Chinese enterprises Alibaba and Tencent. Admittedly there are other metrics of size that includes metrics of revenues, profits, customers and the scope and impact of a corporate enterprise, but the considerable market capitalization of these seven companies place them in the global top ten, which makes them *big*. But are they *bad*? When is an enterprise so big that failure is untenable in terms of social stability?

The global financial crisis of 2008 explored the concept of "too big to fail" in the financial world. Do we have a similar situation with some or all of these digital service enterprises?

At the start of the twentieth century a member of the US Supreme Court, Louis Brandeis, argued that big business was too big to be managed effectively in all cases. He argued that the growth of these very large enterprises that were at the extreme end of the excesses of monopolies, and their behaviours harmed competition, harmed customers and harmed further innovation. He observed that the quality of their products tended to decline, and the prices of their products tended to rise. When large companies can shape their regulatory environment, take advantage of lax regulatory oversight to take on more risk than they can manage, and transfer downside losses onto the taxpayer, we should be very concerned. It is hard to disagree with Brandeis if this outcome is an inevitable consequence of simply being *big*, and given the experiences of the 2008/2009 financial meltdown we could even conclude that Brandeis' observations apply to the financial sector. But do these systemic abuses of public trust in the financial sector translate to concerns in the ICT sector? Brandeis' views did not enjoy universal acclaim. Others at the time, including President Theodore Roosevelt, felt that there were areas where there were legitimate economies of scale, and that large enterprises could achieve higher efficiencies and lower prices to consumers in the production of good and services by virtue of the

volume of production. The evolution of the auto manufacturing industry in the early twentieth century, and the electricity industry both took exotic and highly expensive products and applied massive scale to the production process. The results were products that affordable by many of not all, and the impact on society was truly transformational. The US administration of the day moved to implement regulatory oversight over these corporate behemoths, but not necessarily act to dismantle their monopoly position.

But if the only oversight mechanism is regulation, have we have allowed the major corporate actors in the digital service sector to become too big to regulate? Any company that can set its own rules and then behave in a seemingly reckless fashion is potentially damaging to the large economy and the stability of democracy. One need only mention Facebook and elections in the same sentence to illustrate this risk of apparently reckless behaviour.

To quote Brandeis again: "We believe that no methods of regulation ever have been or can be devised to remove the menace inherent in private monopoly and overwhelming commercial power."

But if we choose to reject Brandeis' view and believe that regulation can provide the necessary protection of public interest, then it is reasonable to advance the proposition that we need to understand the activity we are attempting to regulate. Such an understanding might be elusive. In the digital networking world, we are seeing more and more data traffic go 'dark'. Content service operators are using their own transmission systems or slicing out entire wavelengths from the physical cable plant. This withdrawal of traffic from the shared public communications platform is now not only commonplace, but the limited visibility we have into this activity suggests that even today the private network traffic vastly overwhelms the volume of traffic on the public Internet, and the growth trends in the private data realm also is far greater than growth rates in the public Internet.

How can we understand what might constitute various forms of market abuse, such as dumping, deliberate efforts to distort a market, or discriminatory service provision when we have no real visibility into these dark networks? Yet these dark networks are important. They are driving infrastructure investment, driving innovation and indirectly driving the residual public network service. Are we willing and able to make an adequate case to expose, through various mandatory public filings, reports and measurements, the forms of use of these privately owned and operated facilities and services? Do we have regulatory power to do so considering the size of the entities we are dealing with. We've seen in the past the many national regimes have attempted to avoid the test of relative power by handing the problem to another jurisdiction. The anti-trust action against Microsoft was undertaken in Europe and even then the result was largely unsatisfactory. Even if we might believe that greater public exposure of the traffic carried by the dark networks might be in the public interest we might simply not have the capability to compel these networks operators to undertake such public reporting in any case.

Consolidation

The internet has been constructed using a number of discrete activity areas, and in each area appeared to operate within a framework of competitive discipline. Not only could no single actor claim to have dominate or overwhelming presence across the entire online environment, but even in each activity sector there was no clear monopoly position by any single actor.

Carriage providers did not provide platforms, and platform providers did not provide applications or content. The process of connecting a user to a service involved a number of discrete activities and different providers. The domain name being used can from a name registrar, the DNS lookup was an interaction between DNS resolver application and a DNS server host, the IP address of the service was provided by an address registry, the credentials used for the secured connection came from a domain name certification authority, the connection path provided by a number of carriage providers, and the content was hosted on a content delivery network, used by the content provider. All of this was constructed using standard technologies, mostly, but not exclusively defined by the IETF.

This diversity of the elements of a service is by no means unique, and the telephone service also showed a similar level of diversity. The essential difference was that in telephony the orchestration of all of these elements was performed by the telephone service operator. In the Internet it appears that there is no overarching orchestration of the delivered composite service. It would be tempting to claim that the user is now in control, but this is perhaps overreaching. Orchestration happens through the operations of markets, and it would appear that the market is undertaking the role of resource allocation. However the user does have a distinguished role, in that it is the users' collective preference for services that drives the entire supply side of this activity.

But this is changing, and not necessarily in a good way. Services offered without cost to the user (I hesitate to use the term "free" as this is a classic two-sided market instance where the user is in fact the goods being traded to advertisers) have a major effect on user preferences. However there is also the issue of consolidation of infrastructure services.

As an example, Alphabet not only operates an online advertising platform, but also a search engine, a mail platform, a document store, a cloud service, a public DNS resolver service, a mobile device platform, a browser, mapping services to name just a few. It appears that in this case it is one enterprise with engagement in many discrete activities. The issue with consolidation is whether these activities remain discrete activities or whether they are being consolidated into a single service.

There are two recent examples where this is a likely concern.

The first is the recent specification of DNS resolution over HTTPS (DOH). The DNS is a widely abused service. Attackers often leverage the DNS to misdirect users to the wrong destination and then may attempt various forms of fraud and deception. National content control systems often rely on manipulating DNS responses to make it impossible, or more realistically mildly difficult, to reach certain named service points. The DNS is often used to understand what users are doing, as every Internet transaction starts with a resolution of a name to an address. Observing an individual user's DNS queries may well be enough to profile the user to a reasonably high degree of accuracy. The IETF had its moment of epiphany in the wake of the Snowden disclosures, and undertook a concerted effort to shore up its protocol to prevent casual or even quite determined attempts at eavesdropping. The DNS has been an integral part of this effort and we have seen the specification of DNS over TLS as a way of cloaking the content of DNS queries and responses from observation. DOH looks like a very small change from DNS over TLS, as they both use very similar formats on the wire. However, DOH treats the DNS response as a web object. It can be cached. It can be pre-fetched. Presumably it can be embedded in web pages. This creates the possibility of a browser defining its own DNS environment completely independent of the platform that runs the browser, independant of the local service provider and even independent of the DNS as we know it. If the browser can consolidate name resolutions functions into the operation of the browser itself then it need not rely on a distinct name resolution system, or even a distinct name system. The browser can consolidate names and name services into its own space. Given that some 80% of all user platforms use Chrome as their browser these days then that places a huge amount of unique market power in the hands of the Chrome browser and its provider, Alphabet. DOH may make the DNS a secret to onlookers, but once it's a secret then its beyond conventional oversight and public purview, and whether the consequent deeds in this darkened space are good or bad are effectively impossible to determine.

The second is the use of the QUIC protocol. Applications have normally followed a conventional model of using the underlying operating system for common functions. There are operating system interfaces for working with the local file store, for various network services, such as the DNS and for network connections and the protocol to service the connection, such as TCP. TCP operates with its flow control parameters in the clear, so that network operators may deploy so-called middleware to override the TCP session behaviour and impose its own view of session throughput. It can be a very effective manner of allowing the network operator to discriminate across traffic types, selectively supressing the network demands from less preferred session flows and allowing other sessions to achieve preferred performance. QUIC, originally developed by Alphabet and implemented in Chrome browsers changes all that. Chrome includes its own implementation of an end-to-end flow control protocol within the browser and speaks to its counterpart at the remote end of the connection. The way it does this is to use the IP datagram service (UDP) from the host platform and use an

inner encapsulation to support an end-to-end protocol in precisely the same way that TCP is supported within IP. QUIC also protects itself from observation and manipulation by encrypting its payload. In so doing the browser is consolidating the end-to-end flow control protocol into the browser and not permitting either the host platform's operating system nor the network to have any visibility into the flow state. Like DOH, QUIC drags the end-to-end protocol into a darkened state within the browser.

Both of these are examples of a deeper and perhaps more insidious form of consolidation in the Internet than we've seen to date with various corporate mergers and acquisitions. Here it's not the individual actors that are consolidating and exercising larger market power, but the components within the environment that are consolidating. Much of this is well out of normal regulatory oversight, but the results are not dissimilar to the outcomes of corporate consolidation. The result in these two cases of application consolidation are that the browser provider attains significant gains in market power.

Measurements and Data-Driven Policy

A number of countries, including the United States embarked on consumer broadband measurement projects over the past decade. The effort was intended to remove much of the confusion in the consumer market over the actual differences between broadband access products and allowed the consumer to compared advertised performance claims of a product against recorded experience with the product. A side-effect, intended or otherwise, was that it provided the sometimes-vexed network neutrality debates with some helpful baseline data about whether there was an issue at all of congestion of access networks and to what extent traffic prioritisation mitigated these issues for some content and services and not others.

But once we embark upon such projects it's easy to scan the environment and see many more opportunities where data will help in understanding the market and understanding what regulation, if any, would improve the market both in terms of consumer protection and overall economic efficiency. Today we see various sliced network platforms that take an underlying common infrastructure and present a virtualised view. This space includes the full range of cloud services, APIs, content management, distribution, load shedding, traffic scrubbing and related availability and performance managers of content and services. Questions relating to characteristics of these services, their level of use, their actual performance, their pricing functions and the efficiency of the supply side of these services all reflect a set of public service concerns that are echoes of the original concerns over the broadband access markets. In order to understand that these concerns are at such a level that might require some form of regulatory oversight it is necessary first to define metrics of this sector and then embark on data collection. Of course, any steps in this direction would assuming that we acknowledge that the evolution of the broader supply side of this content and digital service market has reached a level of opacity that its internal structure is not already clearly visible.

The aspects of such problems and concerns include a privacy index, that may entail the structured measurement of privacy in the handling of user data. A consolidation metric could measure the degree to which various common infrastructure services are being consolidated into a small number of actors. Traffic indices could describe the volumes of traffic carried over privately managed, hybrid or public carriage paths.

Online Content Controls

"No matter where you go, there you are!" But how do you know where that is? And how do you get where you want to get to if you don't know where you are?

These days governments are looking to social media platforms to exercise self-moderation. They are coercing the platform operator to take down or otherwise limit the dissemination of content that is considered socially harmful. Proponents of content moderation complain that social media platforms are generally insufficiently responsive to harmful content and the damage it causes. In an effort to address these concerns, companies often publish "transparency reports" detailing the number of complaints received, number of actions taken, and other statistics. Germany's recently passed NetzDG, which requires social media companies to take down content illegal under German law within 24 hours of notification, and to publish a report detailing the number

of complaints, number of takedowns, number of appeals, and number of reinstatements following appeals. The first such report was released in August 2018, and the report clearly shows that social media platforms are taking this measure seriously, but at the same time it highlights a concern that they may be over-zealous in their application of such censorship of content

The democratic processes in our society flourish within a structure of public debate, and that implies that we need to understand the delicate border between what we would consider harmful and what is a controversial. The risk in these operator-managed content control exercises is that they err towards removing all controversial content and influence our social judgement on what we are prepared to consider controversial. The transparency reporting model as used by the NetzDG does not provide any meaningful insight into this dimension of content control being exercised. It is unclear what standards are being used to censor online material. This is an important consideration given the high likelihood that more nation states will emulate the NetzGD model, and rather than passing laws per se about content. If the aim is to make users feel "safe" then takedowns and feelings of safety are not necessarily correlated. It's a complex area and made more so because we find it hard to articulate our actual goals and concerns about the balance between free speech and corrosive and harmful social influence.

Digital Dark Matter

Services without cost are a challenge to economists. Open source technology, including various flavours of Unix, Wikipedia, Yandex and Github are good examples. Google's search and Gmail services can be considered in the same manner. Unlicensed spectrum used by WiFi services fall into the same category. They are inputs into the production of goods and services and can be replicated without limit yet they play no role in the calculation of GDP. They do not contribute to calculations of revenue nor in the calculation of retained wealth.

Ignoring the intrinsic value of these freely provided services can result in skewed data and skewed policies. For example, how can we calculate the true value of public R&D spending in the economy if we ignore the indirectly generated value by the production of open source technology? How effective are our economic policies if the input data ignores the intrinsic value of these unpriced digital goods and services?

Do we observe a more productive IT sector and higher estimate of its value in the national economy where there is a greater level of both the use and the production of these unpriced digital goods. More prosaically, why do some countries make far greater use of open source than others? And is there a correlation between such use and the efficiency and value of the national ICT sector?

It seems that much economic measurement is based on the production and movement of tangible goods and priced services, which may well relate to the structure of national economies in a post-industrial age world. The Internet has introduced the concept of free services, typified by open source technologies, and the challenge we face is to measure the economic value of these forms of digital services and place them into the broader context of national economic measurement in order to formulate effective economic policies.

Universal Service

One of the cornerstones of the deployment of the telephone was the social contract that was expressed in terms of a Universal Service obligation. Private operators were licensed to provide public communications services to consumers, but in so doing they also took on some additional commitments. They were obliged to provide an affordable service to all customers, including rural and remote where feasible, and were not able to only provide service to the richest of customers who were the cheapest to service. This entails some form of structural cross subsidisation in order to ensure that the price of the service were affordable even when the costs were far higher. It also required a clear understanding of what the service was.

In the telephone world the service was essentially the provision of intelligible human voice. However no such underlying model exists in broadband digital services. Should be perform structural cross subsidisation to bring

4K HD streaming video to all? Or do we all need access to services that than sustain a data delivery rate of 10Mbps? Or 100Mbps? In the United States the FCC sets a broadband service at a speed of 25Mbps to the end point and a return speed of 3Mbps. This represents a degree of compromise between what can be cost effectively delivered in dense metro areas, where speeds of 1Gbps and even 10Gbps are present in some markets and what can be delivered in the low density rural and remote environments.

Even when we have a metric of what broadband access means, how can we measure progress for universal service? Detailed maps that can indicated unserved locations are notoriously difficult to assemble and maintain in respect to low density rural and remote landscapes. When such a data collection is assembled how can we coordinate the service provision exercise to ensure that the gaps in service are efficiently filled across the entire set of service gaps, rather than having the service provision sector concentrate on the relatively easy and most cost-effective locations.

Sustaining Internet Measurement

We all believe that good policy relies on good data, but we are not so clear on how to generate good data. In is increasingly challenging for objective observers to undertake measurements, and the collective concern is that the Internet environment is increasingly going dark with respect to third party measurements. If the only available measurements are those undertaken by the service providers themselves then it is challenging to assess the accuracy, completeness and relevance of such measurements as being an objective and complete indicator of the activity.

Are there ways we can provide adequate incentives for measurement regimes? Are we relying solely on cooperation, corporations or confiscation, or are there other mechanisms that would put the measurement activity on a sustainable foundation? Are third party measurement frameworks necessarily funded from the public purse, or can the sector itself develop measurement frameworks that are both funded and undertaken by the sector? We have reached the point where we understand that we need a solid foundation of data to underpin relevant and effective policy regimes, but we are still some distance away from understanding how we can sustain the measurement activity itself.

Two-Sided Markets

Many Internet ventures operate as two-sided markets serving both upstream sources of content and applications as well as downstream consumers. There is a distinct network effect here where sellers prefer to use markets that contain larger number of buyers znd buyers tend to prefer markets with a large number of sellers. These Internet ventures have exploited "winner take all" networking externalities. Regulatory agencies acknowledge the substantial market shares these ventures have acquired, but most refrain from imposing sanctions on grounds that consumers accrue ample and immediate benefits when platform operators use upstream revenues to subsidize downstream services. Consumers also gain when intermediaries eschew short term profits to acquire greater market share. However, it's possible that this tolerance may be wearing thin, and public opinion appears to be turning against these massive platform operators.

A more complete assessment of consumer welfare balances the downstream market enhancements of convenience, cost savings, free-rider opportunities and innovation with upstream aspects of the true value of uncompensated consumer data collection, analysis and sale. Without a rigorous analysis of both upstream and downstream markets and their interaction it is not easy to assess the complete value of these markets, nor is it possible to assess whether the harms of the large-scale platforms that have all but excluded any effective form of competition are outweighed by the extent of benefits realised by consumers.

Observations

It would be rash to describe the Internet as a mature and staid sector. The major characteristics appear to be speed and amplification. Behaviours change rapidly and the Internet appears to amplify many of these changes.

The abuse profile of the Internet has turned much of the networked environment into a toxic wasteland. Any provider of online services who requires assured service delivery has to resort to using one of a very small number of heavily fortified service delivery platforms. It's as if we're returned to 13th century England and security is provided by fortified castles. Is it acceptable that the service delivery space has only a handful of major platform providers? The fact that these protected 'castles' exist as a pragmatic response to a degraded and abused Internet could be seen as a beneficial outcome. Even the fact that the harnessing of resources that allow such castles to be constructed and operated means that there are only a few castle operators is not necessarily harmful in and of itself. The fact that the internet appears to operate efficiently, effectively and far better than previously is beneficial outcome. But what happened to the vision of information technology and the internet as an egalitarian social force that empowered individuals to express themselves directly. This particular vision has been lost, perhaps irretrievably. And what can we say of the few large-scale service platform providers? Is big necessarily bad? It does not seem to be the case, particularly when we observe that big is a necessary precondition to provide services in today's Internet that is well beyond the individual capabilities of smaller actors.

But there are now new barriers to competitive entry. It appears that many new ventures in this space rely on the permissions from one of more of these major service platform providers, and one for of recognition of their success is their acquisition by one of these giants. We need to look no further than GitHub, Skype and WhatsApp. As of December 2018, Alphabet had acquired over 220 companies. One of more effective definitions of true market control is that the incumbents set the terms and conditions of any competitive market entrant, and it certainly appears to be the case that new entrants in this space only thrive with the forbearance of these major incumbents, not by virtue of any public policy of encouragement of competition in these markets.

Perhaps it's time to accept what we might not have accepted in the past and just move on. If the Internet is a world of a handful of giant enterprises, then are we indulging ourselves here if we think that competition-based policy objectives are attainable in any meaningful way? Clinging on to now-ineffectual public policy frameworks is neither reassuring nor productive.

However, in recognising that our digital world is one who's core is populated by a small number of monopolies in their selected service realms does not necessarily imply that there is no problem. Monopoly theory and anti-monopoly laws are based on a view of market stasis where the monopoly resists innovation and disruption in order to entrench its position. Monopolies were conventionally large-scale enterprises with a large labour force, a large commitment of capital, and extensive levels of social influence. The change in today's set of Internet giants lie in the flattening of customer relationship structures where a relatively small enterprise in terms of the size of its labour force and its inventory of capital assets can make extensive use of technology to maintain billions of individual customer relationships. It is this agile grasp of the forefront of technological capability that provides their competitive advantage, and the grip on intellectual property rights that allows it to shrug off various efforts to construct competition. Their services and prices are achieved through the leverage of two-sided markets where bulk in one market drives bulk in the other through these network effects. We have reached the point where these enterprises are not directly replaceable, not substitutable and susceptible to many forms of regulatory intervention without risk of significant social impact. It makes it very difficult to assume that the old remedies will be effective in this new world.

Should we despair? Is public policy in this space now emasculated? Are we experiencing a second generation of the Gilded Age where the new giants get to write the regulations and create the rule set? It's not that there won't be rules and regulations, but the influence that shapes these policies is those of the major incumbents.

But perhaps it's not that bleak. One could make the case that the fact that Alphabet spent USD 18M in 2017 lobbying US politicians shows that Alphabet takes the public policy process very seriously indeed. An optimistic

view is that public policy is still important and is still respected by industry players. The challenge is to find the point of leverage that offers some positive outcomes for the public good inside the cut and thrust of the large policy process. It's not going to be easy. The twitterization of the policy debate, painting complex issues with a simplistic 140-character brush serves no productive purpose. These are complex problems and require considered policy responses.

Is surveillance capitalism destroying the market economy? The market economy relies on substantial capacity in the market, allowing for both suppliers and consumers to have choices. The problem with today's activities is that the anonymity permits price distortions and corrupts the efficient working of the market. Markets took on the role of resource allocation in our economy only two centuries ago and Karl Marx was one of the first to study and describe this. Indeed, he appears to have exceeded Milton Friedman in his enthusiasm for markets as an efficient distributor of resources. If you distort the operation of the market as an allocator of resources, then we undermine social fairness. Such a line of thinking leads to the potential conclusion that GDPR is a small step, and we should go further to compel these user profile information collectors and their users to openly disclose their transactions and trades? Should we consider asserting that users have proprietorial rights to their own personal profile, and should we require individual to grant explicit permission for such transactions?

Measuring consolidation is a difficult problem. As an example, Staples, an office supplies retailer attempted to merge with Office Depot in their efforts to counter their collective revenue loss to Amazon. Such horizontal mergers have a difficult time with the Department of Justice in the US, and the merger proposal was not cleared. The result is of course inevitable, as neither enterprise has the volume to effectively compete with Amazon and both are now in challenging circumstances because of Amazon's overarching presence. On the other hand, it could be argued that the acquisition of WhatsApp by Facebook should not have been approved, as the combination of social media and messaging further consolidates Facebook's position with its user base and makes in a must-use platform for digital advertisers. The arguments for regulatory decisions to allow or block mergers depends on forecasted outcomes, and to conduct usable forecasts we need accurate data.

In times of fundamental change our understanding of the mechanics of the former world just aren't that helpful anymore. A GDP figure that cannot measure the true economic value of freely offered services is not helpful any more. Enterprises now straddle many market sectors and the network effects create fertile incubators that rapidly produce dominant players that assume overarching control within their chosen activity sectors. We navigate our way through this with a public policy framework that attempts to balance the public interest against the self-interest of the private sector. But to have an informed, relevant and effective public policy process we need to understand this changing world. It seems to me that open measurement platforms and open data sets are more important than ever before. We need public measurements that are impartial, accurate, comprehensive and of course unbiased as an essential precondition for the fair and effective operation of markets.

My thanks to the workshop hosts, CAIDA and MIT for inviting me to attend the workshop. It has been a stimulating couple of days for me!

Disclaimer

The above views do not necessarily represent the views or positions of the Asia Pacific Network Information Centre.

Author

Geoff Huston B.Sc., M.Sc., is the Chief Scientist at APNIC, the Regional Internet Registry serving the Asia Pacific region.

www.potaroo.net