Social media research to understand service quality

Ipsos MORI

April 2019

Contents

Executive summary

1 Executive summary

1.1 Background and objectives

Ipsos MORI was commissioned by Ofcom to understand and explore concerns and complaints raised on social media about fixed and mobile telecoms providers. The study was also intended to explore how analysing social media content might inform Ofcom about emerging areas of consumer harm.

The research set out to answer three key questions:

- What concerns and complaints about fixed and mobile telecoms services do consumers in the UK raise and discuss on social media?
- How does the number and nature of complaints and concerns on social media vary between service providers?
- To what extent can analysis of social media be used by Ofcom to identify emerging areas of consumer harm?

Ofcom identified nine fixed and mobile telecoms providers[1] for inclusion in the study, along with Openreach. A total of 121,004 posts relevant to these providers from the full 2018 calendar year were collected from Twitter and Facebook using the social media analytics platform Synthesio[2] based on Boolean search queries developed and run for each provider.

The final clean data set contained 95,241 posts, and was analysed through topic modelling, and factor, sentiment, and qualitative analysis. The range of analysis techniques used allowed for thorough exploration of the data. For example, topic modelling identified the most common themes within the data, while factor analysis grouped these together based on the strength of relationships between different topics. Sentiment analysis identified whether posts contained positive, negative or neutral sentiment, while qualitative analysis allowed for a more detailed and nuanced exploration of different themes and issues within the data.

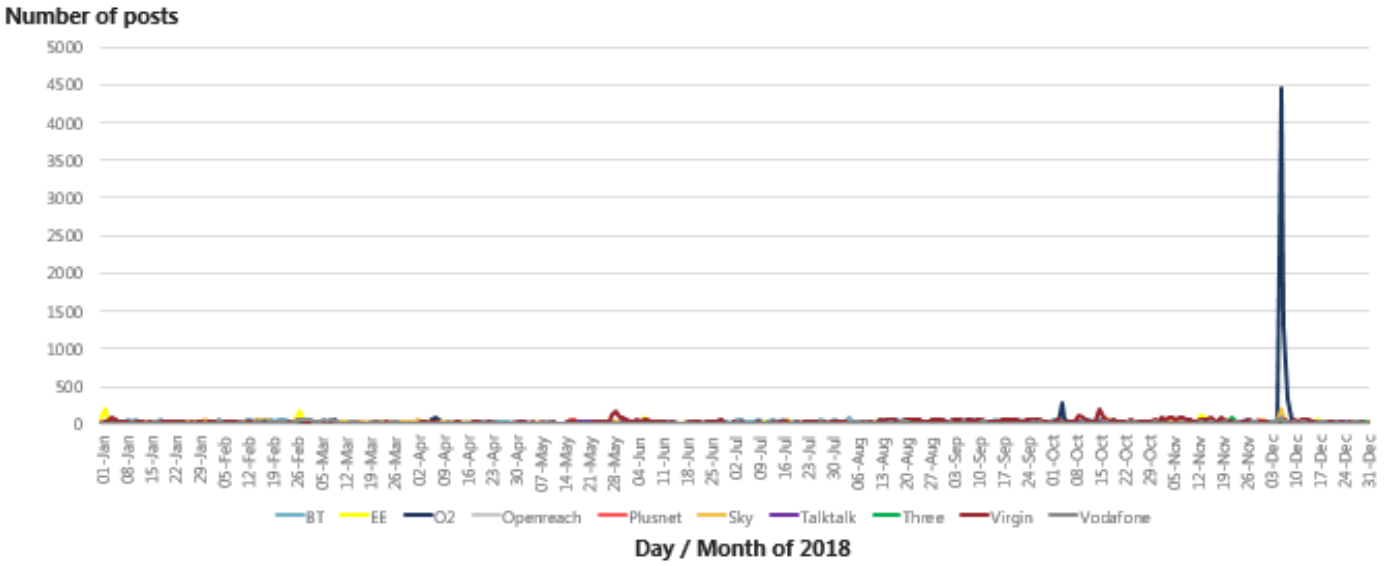
The data set only included publicly available posts from both Twitter and Facebook. Personal identifying information such as Facebook usernames and Twitter handles were not included in the analysis.

1.2 An ongoing conversation, in part driven by events

Consumer posts to and about telecoms providers were ongoing throughout the year, providing rich insight in to the daily experience of telecoms users. This research captured a total of 74,379 posts made by consumers from January 1st through to December 31st 2018, excluding re-tweets; this translates into an average of 204 per day. However, as a source of real-time comment, the dataset also reflects some major events. This is demonstrated by the significant spike in activity on the 6th-7th December, relating almost solely to the O2 network outage. Posts referring to O2 on these days accounted for around ten per cent of the total posts across all providers in 2018. Further spikes in consumer posts related to similar service outage issues for EE, Sky, and Virgin Media, as well as downtime in customer services platforms such as websites and chatbots.

[1] BT, EE, O2, Plusnet, Sky, TalkTalk, Three, Virgin Media, Vodafone

2 https://www.synthesio.com/



1.3 Online conversations offer rich insights into the nature and tone of consumer concerns

Initial analysis of the unstructured dataset (including provider and consumer posts and re-tweets) used topic modelling to identify patterns within the language used in social media posts.[3] The resulting topic model identified 16 key topics, underpinned by 39 subtopics. The breadth of topics identified within the data demonstrates that no single topic dominates online conversations about fixed and mobile telecoms providers.

Table 1.1: Topic and subtopic counts, including retweets, consumer, and provider posts

Topic	Subtopic	Facebook	Twitter	Total
Time	Hours, Days, Weeks	21%	17%	19%
Tone	Consumer - Rage	5%	3%	4%
	Consumer - Escalate	5%	2%	3%
	Consumer - Undervalued	4%	1%	3%
	Provider - Apologetic	13%	3%	8%
	Provider - Diligent	2%	<1%	1%
	Provider - Sympathetic	2%	<1%	1%
	Tone total	26%	9%	17%
Cost	Bills	18%	9%	13%
	Price changes	2%	2%	2%
	Payment method	1%	<1%	1%
	Cost total	19%	11%	15%
Communication method	Web chat	9%	5%	7%
	Private message	9%	1%	5%
	Email	5%	4%	4%

[3] Topic modelling uses a number of stages to identify the topics in a large, unstructured dataset – using a combination of algorithms and human review. The approach used here involved applying tags to posts based on specific terms and parts of words or phrases which can be grouped together. Lists of similar and related terms are then created. Finally, non-negative matrix factorisation was used to combine terms and attributes to create topics and themes. Further details of the approach are included in Chapter 3.

	Phone	1%	1%	1%
	Communication method total	**21%**	**10%**	**15%**
Poor customer service	Customer service	**20%**	**10%**	**15%**
Solutions	Acknowledgement	9%	4%	6%
	Reset	6%	4%	5%
	Diagnose	1%	2%	1%
	Guidance	3%	1%	2%
	Check	3%	1%	2%
	Solutions total	**17%**	**11%**	**14%**
Signal and coverage	3G and 4G coverage	8%	5%	6%
	Network outage	1%	2%	2%
	Signal and coverage total	**9%**	**7%**	**8%**
Hardware	Street level	3%	4%	4%
	In home	2%	3%	2%
	Sim cards	2%	1%	1%
	Hardware total	**7%**	**8%**	**7%**
Wifi	Reach	3%	3%	3%
	Connection	2%	1%	1%
	Wi-fi	2%	1%	2%
	Wifi total	**7%**	**5%**	**6%**
Speed	Speed	**6%**	**6%**	**6%**
Engineers	Engineers	**6%**	**4%**	**5%**
Contract	Renewal and end	5%	2%	3%
	Early termination and charges	<1%	<1%	<1%
	Cancellation period	<1%	<1%	<1%
	Transferring provider	<1%	<1%	<1%
	Contract total	**6%**	**2%**	**4%**
Data	Usage and allowance	**3%**	**3%**	**3%**
Account	Details	**5%**	**1%**	**3%**
Broadband	Broadband	**3%**	**3%**	**3%**
Consumer impact	TV and film viewing	1%	1%	1%
	Emergencies	1%	<1%	<1%
	Consumer impact total	**2%**	**1%**	**1%**
Uncategorised		**22%**	**36%**	**29%**

Many posts contained two specific types of language, categorised in the topic model as 'Tone', and references to 'Time'. 'Time' posts specified lengths of time such as minutes, hours, days or weeks. Consumers were often frustrated about wait times or the length of time they had been affected by particular issues.

Consumer posts categorised into the 'Tone' topic fall into three subtopics: 'Rage', 'Escalate', and 'Undervalued'. 'Rage' posts tended to be highly emotive, with consumers expressing disappointment and anger. 'Escalate' posts were more likely to include references to official bodies such as 'ombudsman' and 'Ofcom'. 'Undervalued' posts contained language which tended to reflect a sense of being let down, with consumers often referring to themselves as 'loyal customers'.

Provider-driven 'Tone' posts were grouped into 'Apologetic', 'Sympathetic' and 'Diligent' subtopics. 'Apologetic' posts were the most common, using terms such as 'sorry' and 'disappointed'. 'Sympathetic' posts contained terms such as

'understand', 'important' and 'appreciate', alongside adjectives such as 'completely' and 'totally'. The language in 'Diligent' posts included reassuring terms like 'assure' and 'work hard' alongside words suggesting immediacy, such as 'soon' and 'quick'.

1.3.2 What concerns and complaints do consumers raise on social media?

When analysing consumer-only data (excluding retweets and posts identified as being made by providers) topic frequencies remain broadly similar to the full data set. High levels of 'Tone' and 'Time' posts remain, reflecting the way consumers often frame their concerns and complaints.

Table 1.2: Topics excluding retweets and provider content

Topic	Facebook	Twitter	Total
Time	29%	20%	24%
Cost	25%	12%	17%
Customer service	23%	10%	16%
Tone	20%	8%	13%
Solutions	11%	8%	9%
Signal and coverage	9%	6%	8%
Hardware	8%	9%	8%
Communication method	10%	5%	7%
Wifi	6%	5%	6%
Speed	6%	7%	6%
Engineers	7%	3%	5%
Contract	8%	2%	5%
Data	3%	3%	3%
Broadband	3%	4%	4%
Account	2%	1%	2%
Consumer Impact	2%	2%	2%

Posts about 'Cost' were the most frequently mentioned within consumer-only data, with posts mainly related to bills. 'Customer service' posts tended to contain descriptive terms such as 'poor' and 'rude' as well as words relating to conversations such as 'speak' and 'tell'. For the 'Solutions' topic most consumer posts used terms like 'reset', rather than those in the overall topic model that are more likely to be used by providers (e.g. 'guidance' or 'diagnose'). 'Signal and coverage' tended to focus on concerns about 3G and 4G. 'Hardware' posts mentioned sim cards, street level terms (like 'street', 'copper' and 'exchange') and in-home hardware (such as 'plugs', and 'flashing lights' on broadband hubs).

The longer character limit for Facebook posts means they tend to contain richer, more varied, content. As a result of this, a single post on Facebook was more likely to be categorised into multiple topics, and therefore overall topic frequencies are higher for Facebook when compared to Twitter. Allowing for that caveat, there were some notable differences between Facebook and Twitter posts that may be worth exploring further. 'Customer service', 'Costs', 'Contracts' and 'Engineers', were topics which were all more likely to be raised on Facebook when compared with Twitter, perhaps reflecting the different ways consumers use different social media platforms.

1.3.3 How do providers respond to consumers?

Looking at the full data set, excluding retweets, providers generally communicate using apologetic language, and look to suggest solutions, or invite further communication through direct one-to-one channels. Provider posts referred to a range of common 'Communication methods' including 'Webchat', 'Private message', 'Email' and 'Phone'. 'Webchat' and 'Private message' were the most frequently mentioned, suggesting a preference for online communication.

In some instances, providers sought to suggest 'Solutions' to issues in response to consumer complaints. These fell into five categories; 'Acknowledgement', 'Reset', 'Check', 'Diagnose' and 'Guidance'.

1.4 Customer service is the most commonly cited issue online

Building on the initial topic modelling, factor analysis identified ten key issues of discussion within social media data,[4] the most common of which was 'Poor customer service'.

[4] Factor analysis groups together related topics from the topic model, based on the likelihood that they occur near each other within individual posts. The findings are based on all the posts included in the topic model, including provider posts and retweets. This is a useful way to summarise the topic model and begin to explore the connections between issues in posts by consumers and providers.

Table 1.3: Factor analysis to group topics

Factor	Topic/subtopic
Poor customer service (33%)	Time
	Customer service
	Tone: consumer - Rage
	Tone: consumer – Escalate
	Tone: consumer – Undervalued
	Consumer impacts: Emergencies
Costs and contracts (17%)	Cost: Bills
	Contract: Renewal and end
	Costs: Price changes
	Cost: Payment method
	Contract: Cancellation period
	Contract: Transferring provider
	Contract: Early termination and charges
Sympathetic provider responses - suggesting solutions (14%)	Solutions: Acknowledgement
	Solutions: Check
	Tone: provider – Diligent
Apologetic provider responses - inviting private communication (11%)	Tone: provider – Apologetic
	Communication method: Private message
	Account
	Communication method: Email
	Tone: provider – Sympathetic
Broadband and speed (11%)	Speed
	Hardware: Street level
	Broadband
	Consumer impacts: TV and film viewing
Signal, coverage and data (11%)	Signal and coverage: 3G and 4G
	Data: Usage and allowance
	Hardware: Sim cards
	Signal and coverage: Network outage
Contact details (8%)	Communication method: Webchat
	Communication method: Phone
Wifi (8%)	Solutions: Reset
	Wifi: Connection
	Wifi
Engineers and diagnosing issues (7%)	Engineers
	Solutions: Guidance
	Solutions: Diagnose
Hardware (5%)	Wifi: Reach
	Hardware: In home

The first factor, 'Poor customer service', demonstrates how issues relating to customer service are particularly emotive. All three of the consumer tones, 'Rage', 'Escalate', and 'Undervalued' were most likely to occur alongside comments about customer service, demonstrating the strength of feeling about this issue. This factor is also the most likely to contain mentions of Ofcom.

The factor analysis resulted in two provider-driven factors. Both reflected providers offering sympathetic responses, coupled with invitations for further private communication or suggestions for immediate solutions to consumer issues. Invitations for further communication were offered in the form of a 'Private message' or 'Email', alongside a request for more details about consumers' 'Accounts'. Those suggesting 'Solutions' are most likely to suggest that customer perform 'Checks', while the providers language focuses on acknowledging the issue raised and reassuring the consumer that it will be addressed.

All 'Cost' and 'Contract' subtopics were grouped into one factor, reflecting the intertwined nature of these issues. Within this group, the subtopics 'Early termination and charge' and 'Renewal and end' were the most closely related, suggesting that cost and contract issues may be particularly likely to arise at specific times in the customer journey.

The broadband and speed factor illustrated how conversations about speed are often linked to impact on consumers.

Internet and mobile hardware issues are frequently discussed. Direct references to 'Wifi' tend to occur alongside mentions of 'Connection'. Conversations about 'Wifi' are also the most likely to result in in providers suggesting 'Reset' solutions. 'Signal and coverage' and 'Data' topics were the most closely related topics in this factor, illustrating the interactions between topics predominantly related to mobile service.

Mentions of 'Engineers' are often accompanied by provider responses, either suggesting further 'Guidance' or 'Diagnose' solutions.

1.5 There are differences by provider, but explaining these is complex

1.5.1 Interpreting provider differences

At this stage, a direct, robust comparison of comments by providers is challenging. There are a number of interlinked reasons that could help explain some of the differences in findings between different providers, including differences in the size and range of services providers offer, customer demographics, how they manage their online presence, and how the data capture approach allows for different provider names. However, although there are limitations to how social media data can be used to directly compare performance on key issues, indicative comparisons do shed light on the relative nature of the online conversation related to each provider.

1.5.2 Key factors by provider

A useful point of focus is the seven factors from the overall factor analysis that relate most directly to consumer experience. These indicate where it is relatively more or less common for consumers to discuss certain issues about different providers online.

Table 1.4: Overall factors based on consumer data excluding retweets, by provider

Factor	BT	EE	O2	Openreach	Plusnet	Sky	TalkTalk	Three	Virgin Media	Vodafone	Total
Customer service	40%	33%	30%	30%	**50%**	34%	46%	34%	37%	47%	38%
Costs and contracts	20%	21%	18%	10%	**26%**	18%	23%	21%	15%	24%	20%
Broadband and speed	20%	11%	2%	46%	16%	**19%**	18%	7%	8%	11%	13%
Signal, coverage and data	4%	25%	25%	6%	4%	5%	4%	**30%**	5%	16%	11%
Wifi	7%	7%	6%	3%	**9%**	5%	8%	7%	7%	6%	7%
Hardware	10%	4%	1%	**14%**	6%	7%	8%	3%	6%	3%	6%
Engineers and diagnosing issues	9%	2%	1%	9%	8%	6%	**10%**	1%	5%	3%	5%

'Poor customer service' was the most common factor for all providers. Similarly, 'Costs and contracts' issues were raised by consumers across providers. This suggests there are some consistent reasons for consumers to contact providers or discuss issues using social media. Some factors are service specific, related either to mobile or fixed services. These include the 'Broadband and speed' factor, and the 'Signal, coverage and data' factor which were relatively important for fixed and mobile providers respectively.

The 'Signal, coverage and data' factor may have the potential to be refined to better reflect actual performance among mobile providers. Based on the time series data, O2 and EE both had significant outages during 2018, with big spikes of complaints and concerns, and this factor is more common in the online conversation about these providers. This is a smaller issue in the online conversation about Vodafone. However, it is interesting to note that Three had the highest proportion of posts relating to 'Signal, coverage and data', despite not having a major outage, reflecting how this issue was key to consumers beyond specific outage events.

Similarly, the 'Hardware' and 'Engineers and diagnosing issues' factors were more common in the online conversation about providers of fixed services. The latter factor was more common for TalkTalk, Openreach and BT. In the case of the 'Broadband and speed' factor, posts about Openreach were largely driven consumers asking or complaining about superfast broadband not being available in their area.

1.5.3 Key topics for providers

Looking at the underlying topics in the topic model, 'Customer service' was the most common consumer-specific topic. The most frequently mentioned topics for a specific provider were often closely aligned to the providers' services. For example, 'Signal and coverage' was more common for providers who focus more on mobile services.

There were differences in the extent to which providers communicated with their customers publicly on Facebook and Twitter, indicated by the 'Communication methods' and 'Solutions' provider-specific topics. EE, Sky, Virgin Media and Vodafone all had these topic in their top three (excluding 'Tone' and 'Time' topics which are more general), suggesting they may be more likely to respond to consumers on social media than other providers.

2 Background and objectives

2.1 Background to the research

The internet is central to the everyday lives of people in the UK. The last decade has seen rapid development and diversification of online services. Most UK adults are now online and two thirds have some form of social media account.[5] As the communications regulator, Ofcom's role involves overseeing internet service provision offered by mobile and fixed telecoms providers, to promote competition, maintain standards and ensure consumers are protected from bad practices.

Ofcom regularly conducts research to gauge consumer views and experiences of fixed and mobile telecoms services. More recently, social media has become an accessible – and often preferred – channel for consumers to discuss their experiences of services and to contact service providers. Ofcom would like to explore social media research as a methodology that could support their work on behalf of consumers.

Ipsos MORI was commissioned by Ofcom to understand and explore concerns and complaints raised on social media about fixed and mobile telecoms providers. The study was also intended to explore how analysing social media content might inform Ofcom about emerging areas of consumer harm.

2.2 Research objectives

The research set out to answer three key questions:

What concerns and complaints about fixed and mobile telecoms services do consumers in the UK raise and discuss on social media?

How does the number and nature of complaints and concerns on social media vary between service providers?

To what extent can analysis of social media be used by Ofcom to identify emerging areas of consumer harm?

The findings of the research will be used to support Ofcom's annual report 'Comparing Service Quality'[6] and provide further insight into the value of social media research for future work.

[5] https://www.ofcom.org.uk/__data/assets/pdf_file/0011/113222/Adults-Media-Use-and-Attitudes-Report-2018.pdf

[6] https://www.ofcom.org.uk/__data/assets/pdf_file/0022/145525/comparing-service-quality-2019.pdf

3 Methodology

3.1 Overview

Ofcom identified nine fixed and mobile telecoms providers for inclusion in the study, along with Openreach. A total of 121,004 posts relevant to these providers was collected from Twitter and Facebook using the social media analytics platform Synthesio.[7] Posts were collected for the full 2018 calendar year. To collect the data, individual Boolean search queries were developed and run for each provider, with the resulting posts combined for cleaning and analysis. The cleaning process further removed non-UK based or other incoherent content. The final clean data set contained 95,241 posts, and was analysed through topic modelling, and factor, sentiment, and qualitative analysis. The resulting topic model contained 16 key topics, underpinned by 39 subtopics.

This chapter provides further detail of the sources, tools, techniques and processes used to collect and analyse the data. Full descriptions of each provider query can be found in the appendix.

3.2 Data scoping

3.2.1 Social media analysis tools

Ipsos MORI used social media analytics platform Synthesio to collect social media data. Synthesio was chosen as the most appropriate platform because of the size of the database and initial scoping of different platforms demonstrated that Synthesio had a strong coverage of telecoms within the UK. Additionally, Synthesio has a specialised anonymization tool for exporting data. This tool allowed for data to be exported without data fields containing personal identifying information, such as Facebook username or Twitter handle – as such, these personal identifiers were not included in the analysis.

3.2.2 Data sources and coverage

Relevant social media data was collected in relation to nine telecoms providers, specified by Ofcom:

- BT
- EE
- O2
- Plusnet
- Sky
- TalkTalk
- Three
- Virgin Media
- Vodafone

In addition to the nine providers listed above, Openreach was also included in the analysis. Openreach does not offer services directly to consumers, instead working with telecoms providers as its customers. However, end consumers may come into contact with Openreach, for example, when repairs have been organised by their provider, or when raising queries about broadband speeds in their area. Conversations about Openreach on social media were therefore included in the research, as they may form part of the consumer experience.

[7] https://www.synthesio.com/

Posts were collected from Twitter and Facebook for a one year period (01/01/2018 – 31/12/2018). Throughout the report, the term 'posts' is used to refer to individual posts, including comments on Facebook and tweets and replies on Twitter. This means that the figures presented are not based on individual consumers, as some will have multiple posts captured within the data set. As such, any percentages quoted are based on the number of posts, not on the number of individual consumers posting.

It should be noted that the way data was collected from Twitter and Facebook is different, based on the access to data made available by each platform. Within Twitter, the entire public corpus for 2018 was searched on the basis of the queries developed for each provider. This included posts that were made directly to providers, posts that were made about providers, and posts made by providers. By contrast, raw Facebook data was only available to be collected from the public Facebook pages for each of the providers. This therefore includes posts that were made directly on the public pages of each provider and posts made directly from providers to these pages, but not posts that discussed providers in private elsewhere on Facebook. The following public provider Facebook pages were used for data collection:

https://www.facebook.com/O2uk/
https://www.facebook.com/EE/
https://www.facebook.com/Virginmedia/
https://www.facebook.com/Sky/
https://www.facebook.com/BTUK/
https://www.facebook.com/VodafoneUK/
https://www.facebook.com/TalkTalk/
https://www.facebook.com/ThreeUK/
https://www.facebook.com/Plusnet/
https://www.facebook.com/Weareopenreach/

Synthesio accessed Twitter data through the PowerTrack API and Facebook data was collected through the public Facebook API. An API is an 'Application Programming Interface' which allows different types of software to communicate and exchange data, in this case Synthesio, Facebook and Twitter. Only public content was captured from both social media platforms. For Twitter, this included capturing retweets, quotes and mentions. For Facebook this included both comments and replies.

Both consumer and provider responses were collected as part of the same data set. As data was anonymised (author ID was removed), posts were tagged as either consumer or provider on the basis of a machine learning algorithm trained on a manually tagged sample of post. The implications of this approach to tagging are discussed in detail later in this chapter.

Retweets were included in the initial data set to ensure the full breadth of the public online conversation around concerns and complaints about fixed and mobile telecoms providers was captured. This approach allowed for analysis of the data with and without retweets, enabling us to explore whether frequencies for certain topics were noticeably different as a result. However, most of the analysis was conducted *excluding* retweets. This was predominantly because of the challenges around inferring consumers' reasons for retweeting a post, particularly that retweeting a post does not necessarily imply agreement with the content in that post. Additionally, retweets formed a relatively small proportion of the data set meaning the exclusion of retweets did not fundamentally alter the findings emerging from the analysis.

All findings in the report make it clear whether retweets are included or excluded. In the final clean data set, there were c.5,000 retweets. This made up 5% of the total number of posts and 10% of Twitter posts (when retweets were included in the total).

3.3 Differences between data sources

When conducting this type of analysis, it is important to recognise the differences between posts on Twitter and Facebook. Firstly, posts on Twitter are limited to 280 characters, while Facebook posts have a much higher maximum limit of 63,206 characters. This means Twitter content is, on average, much shorter and less detailed, referring to fewer issues

within a single post. By comparison, Facebook posts often contain greater complexity, with more detailed context and varied language. The longer character limit allows consumers to have space to discuss multiple issues within a single Facebook post. It is also important to note that Facebook data was only collected from posts on public provider pages, meaning that private consumer posts could not be included. However, the different nature of Twitter posts allowed for the inclusion of data that was not directly addressed to provider accounts. Tweets containing provider names, but not necessarily official provider handles, were included. This process is discussed in more detail in the following 'query development' section.

There are also differences in the number of users and their demographic profile for Twitter and Facebook in the UK. Two thirds of UK adults say they have a social media account, of these nine in ten say they have a Facebook account, while a quarter say they have a Twitter account.[8] In terms of the demographic characteristics of users, women and C2 households are more likely to have Facebook accounts, while 16-24 year olds and AB adults are more likely to have Twitter accounts.

3.4 Data specification

3.4.1 Query development

Relevant social media data was identified and collected through user defined search queries developed by Ipsos MORI. A 'query' is a search formula that uses a combination of keywords (which are not case sensitive) and Boolean operators (AND, OR, NOT, NEAR) to isolate information. The full queries can be found in the appendices.

Desk research was used to develop the initial queries. This included searching for provider Twitter handles (official and incorrect, as posts are often misdirected to unofficial or wrong pages), as well as common alternatives and misspellings and key broadband and wireless terms. Pilot queries were previewed within the Synthesio platform (which did not require data to be downloaded) and underwent iterative human review. Each individual term was tested and revised until previews of the query displayed content with 80-90% relevancy.

Table 3.1: Summary of how each query was structured

	Consistent across all queries	Varied between queries
Broadband and wireless terms	√	
Provider specific terms		√
Provider products		√
Exclusions - general	√	
Exclusions – provider specific		√

This approach was taken to achieve high levels of consistency and relevancy between queries. Consistency and relevancy were essential for allowing broad comparability between providers and to meet the required standard for analysis, without the need for excessive cleaning. Individual queries were designed for each provider, to allow for provider specific inclusions and exclusions to be made beyond the core, consistent query.

- Broadband and wireless terms: These terms were consistent across all provider queries. These included general terms relating to broadband and wireless services, such as 'Wifi', 'broadband' and 'internet'. A 'bottom-up' approach was taken when deciding which terms should be included in this section. This approached involved using a smaller group of general, relevant, terms to capture a broader set of data. The benefit of a bottom-up approach is that it allows for analysis of a wider data set, using different analytical techniques to draw out themes. The

[8] https://www.ofcom.org.uk/__data/assets/pdf_file/0011/113222/Adults-Media-Use-and-Attitudes-Report-2018.pdf

alternative, which was tested during the early stages of query development, would have been a 'top-down' approach. This involves the inclusion of complaint terms derived from desk research and existing telecoms complaints data. For example, terms like 'speed' or 'coverage'. A top-down approach applies qualitative analysis at the early query development stage, making assumptions about what is expected to be found within the data and therefore creating narrow search terms to target terms covered by these assumptions. This would have risked systematically excluding important aspects of the online conversation about telecoms concerns and complaints.

- Provider specific terms: These included provider names, common variations and misspellings, official and incorrect Twitter handles. Provider Twitter handles for business products and services were also included, where applicable, as desk research highlighted that non-business consumers often incorrectly direct posts to these accounts.

- Provider specific products: These were also added to the queries, and varied by provider, to capture conversations about provider products relating to fixed and mobile services.

- Exclusions - general: A further part of each query consisted of exclusions – meaning posts containing these terms would not be captured in the data set. Each query contained at set of consistent exclusion terms, aimed at removing marketing content and new stories, for example terms such as 'win', 'deal' and 'offer'. Additional terms in this group of consistent exclusions related to irrelevant content frequently flagged during the testing stage. For example, 'dogs' being mentioned alongside the broadband and wireless term 'service' (service dogs) and conversations about the 'internet of things'[9].

- Exclusions – provider specific: The second part of the exclusion terms were provider specific. This was used to exclude unrelated provider services (for example Virgin trains or Sky Sports) as well as irrelevant content associated with provider names. For example, 'level 3' being associated with Three, or 'night sky' associated with Sky. Again, these terms were identified through desk research and extensive query testing and refinement.

3.4.2 Data collection

Using the Synthesio platform, data was filtered on language (English only), location (UK only), the date and time posts were made (01/01/2018 – 31/12/2018 only) and social media platform (Facebook or Twitter only) before downloading. This was to ensure data relevant to the research objectives (UK data over a one year period) was captured.

In addition to the content of each post, a specific set of metadata was captured and downloaded. These fields were:

- ID (unique anonymised numerical identification number attributed to each post)
- Date and time of the post
- Content (text within the Facebook or Twitter post)
- Site (Twitter or Facebook)
- Country
- State (England, Scotland, Wales, Northern Ireland)

Metadata was limited to the fields expected to be useful for analysis based on the research objectives, and to meet the requirements outlined in the project specific data protection impact assessment (DPIA). Personal identifying data was not deliberately collected, including Facebook user name or Twitter handle. The implications of this for analysis are outlined in more detail later in this chapter.

An automatically generated 'interaction count' (number of likes, comments, counts, retweets, shares, mentions and quotes) and 'engagement rate' (rating generated by a calculation of responses to a given post)[10] was included in the

[9] https://data.london.gov.uk/blog/the-trouble-with-the-internet-of-things/

[10] Engagement rate was calculated on Twitter posts as likes + retweets/total potential impressions. For Facebook posts it was calculated as reactions (including likes) + comments/total potential impressions (the 'total potential impressions' is the number of users who could in theory view the post).

download. This was not used in the analysis described in this report as it was found not to be helpful in addressing the research objectives.

The total number of posts were within processing capabilities, and therefore did not require a random sample to be extracted. After the data from each provider query was downloaded, it was manually combined to form one overall data set. The total number of posts included in the downloaded data was 121,004.

3.5 Data cleaning and analysis

3.5.1 Additional data cleaning

Standard machine learning algorithms were used to further clean the data set. This involved manually tagging 1,000 posts for relevance then applying a statistical model to recognise relevant posts. The statistical model applied a 'universal sentence encoder[11]' to convert texts within posts to vectors. These vectors were then used by a 'neural network model', to identify whether posts were relevant.

Text was then cleaned to remove internet links and email addresses, standardise language (e.g. changing 'it's' to 'it is', remove extraneous characters, and resolve spelling mistakes.

The language and location filters applied prior to download do not remove all irrelevant data, so further cleaning was applied to ensure non-English, non-UK posts were removed. A manual quality check was also applied to a subset of the data (1,000 posts) to ensure relevancy was maintained.

After further cleaning stages the total number of posts in the data set was 95,241. Exclusions mainly related to non-English and non-UK posts not previously detected, and some posts which were originally public, but had since been deleted or made private.

3.5.2 Data analysis

The data set was analysed in a number of ways: using topic modelling, factor analysis, automated sentiment analysis, and human qualitative analysis. These types of analysis are discussed in more detail below.

Analysis does not include emoticons or 'emojis', with the exception of sentiment analysis (where these are included as part of the standard algorithms).

1) Topic modelling and factor analysis

Data was analysed using the Ipsos MORI in-house topic modelling platform, built in Python. This applied the 'Spacy English model'[12] and used name entity recognition and lemmatisation. This involves applying tags to posts based on specific terms (such as provider names) and parts of words or phrases which can be grouped together as one (for example, 'mobile' and 'mobiles', or 'TV' and 'television'). 'Key term extraction' was then applied, which involved creating lists of similar and related terms. This included telecom providers, telecom services and hardware, and positive and negative comments.

Different approaches were then applied to the data and it was found that 'non-negative matrix factorisation'[13] was the most appropriate topic model method for the data set, as it resulted in meaningful topics which could be clearly

[11] For further details please see: https://tfhub.dev/google/universal-sentence-encoder/2

[12] https://spacy.io/usage/linguistic-features/

[13] https://docs.oracle.com/database/121/DMCON/GUID-76F89641-E1D3-4B11-8319-4A152389D510.htm#DMCON349

interpreted[14]. Non-negative matrix factorisation analyses and combines terms and attributes to create topics and themes. This method is particularly useful when dealing with data sets containing large varieties of terms which may be ambiguous or initially difficult to interrelate.

The outcome of this analysis is a topic model – a model reflecting different topics within the data, broken down by provider, site (Facebook/Twitter) and sentiment (positive, negative and neutral). A single post can be allocated to multiple topics, meaning totals do not always sum to 100%. Initially 46 topics were identified, before being qualitatively reviewed, refined and relabelled to create 39 subtopics categorised into 16 overarching topics.

Of the total clean data set (95,241 posts), 71% was classified into topics within the topic model. The remaining 29% of data from the data set was not included due to posts being irrelevant, or forming topics which were too small. This is described in more detail in the 'Overview of the conversation' chapter.

The data included in the topic model was analysed for mentions of Ofcom. A total of 737 posts contained the word 'Ofcom', which reduced to 610 after retweets and provider responses were excluded.

A factor analysis was also used. This involved statistical analysis to map relationships between topics and group them based on terms commonly occurring next to or near each other within a post. Strongly related terms and topics were grouped together, forming a total of 10 factors.

2) Sentiment analysis

A sentiment is the indicator that determines whether a post contains positive, negative, or neutral language. This is assigned automatically in this research by a pattern.en algorithm[15] which analyses the number of positive or negative words and emojis in a post, based on a predetermined list (e.g. good, great, amazing etc). Each part of a post is analysed for sentiment meaning a single post can be classified multiple times. For example, a post which contains a sentence complaining about terrible internet connection, followed by a sentence which talks about great customer service, may be categorised under both the positive and the negative sentiment category. This is particularly important to acknowledge for Facebook data where there is an increased likelihood that longer posts will contain different sentiments. Sentiment analysis was run on topics and factors.

3) Qualitative analysis

Qualitative review and analysis was applied throughout the research, particularly for refining the topic modelling and factor analysis, by experienced Ipsos MORI researchers. Qualitative review was essential to further unpick the themes identified by algorithms, and interpret the meaning and nuances within the data. This kind of qualitative analysis, by its interpretive nature, involves a degree of subjectivity. Ipsos researchers interpreted the data, framing their analysis with reference to the research objectives and background to the research. This involved making judgements about which topics and factors to include.

3.6 Data collection considerations and limitations

3.6.1 Representivity

Specific considerations need to be taken into account when considering sampling for social media research. Social media platforms are, by their nature, self-selecting and therefore the data collected from these platforms will not be representative of the UK population, nor will it be representative of all consumers for each provider. Further research is

[14] Non-negative matrix factorisation, latent dirichlet allocation, and factorisation on embedded text similarities were different approaches tested on the data set as part of the topic modelling process, before identifying non-negative matrix factorisation as the most appropriate.

[15] https://www.clips.uantwerpen.be/pages/pattern-en#sentiment

needed to consider social media data in the wider context of consumer complaints, to help identify which consumers are using social media to discuss or to complain to providers, and why.

The data analysed in this report represents the number of posts, and does not represent the number of individual consumers within the conversation. It is likely that some issues elicit multiple contributions or responses, where others may be just as important to the consumer but are mentioned less frequently. It is also likely that a small number of consumers are more prevalent within the dataset, primarily because they are more active users of social media. The data is therefore not suitable for direct comparison to survey research, where each respondent only counts once, or in the context of a multi-coded question, where each issue can only be mentioned once by each respondent.

3.6.2 Differentiating between fixed and mobile related complains and concerns

Concerns and complaints raised on social media tend not to explicitly refer to fixed or mobile services. This is more likely in Twitter data due, in part, to the restricted character limit. The variety of analysis methods applied seek to unpick these differences to some extent. For example, through the creation of topics containing terms such as 'Wifi' and 'broadband' and qualitative analysis. However, this difficulty is further compounded because providers offer different portfolios of both fixed and mobile services. Therefore, due to the difficulty in definitively and consistently identifying the relevant differences across the data set, findings do not focus on differences in the online conversation for fixed vs. mobile services.

3.6.3 Anonymity

It is technically possible to export the author associated with each post as part of the dataset collated by Synthesio and other social media aggregated tools. However, with a view to ethical and GDPR requirements, it was decided not to export the 'author' field within the data analysed by Ipsos MORI. There were a number of implications this had for the analysis, and these are referenced throughout the report.

One challenge is around clearly differentiating between consumer and provider posts, when classifying information is not included in the data set. To mitigate this, an algorithm was created to identify provider posts based on a manually tagged sample. In total, 1,000 posts were tagged and the higher confidence interval of 0.75 (over 0.5) was used to allocate posts as either provider or consumer posts. The majority of posts were made by consumers (c.75,000 excluding retweets) compared to providers (c.16,000 excluding retweets). As it is not possible to definitively conclude that all posts labelled as provider posts were made by providers (even with a comprehensive human review), it is important to take this into account when considering the findings.

Total posts, after cleaning, broken down by consumer and provider content:

Table 3.2: Total posts, after cleaning, broken down by consumer and provider content

	Consumer posts (n)	Provider posts (n)	Total posts (n)
Including retweets	79,240	16,001	95,241
Excluding retweets	74,379	15,859	90,238

While author was excluded from the data set, it is important to note that some incidental personal information may have been captured. For example, a consumer referring to the street they live on when complaining about their network speed. However, instances where this data may have been inadvertently captured have not been published.

Finally, it should be noted that within this report, anonymity has been maintained by 'masking' quotes. This involves making slight changes to posts to ensure they cannot be reverse searched on the relevant social media platform to identify the original user, without significantly changing the meaning of the post.

4 Overview of the conversation

This chapter outlines the overall findings from analysis of Twitter and Facebook posts captured through the provider queries, as described by the topic model. The focus is on the main topics identified in the online conversation overall, and what that tells us about concerns and complaints when it comes to fixed and mobile telecoms providers.

4.1 Summarising the conversation – the topic model

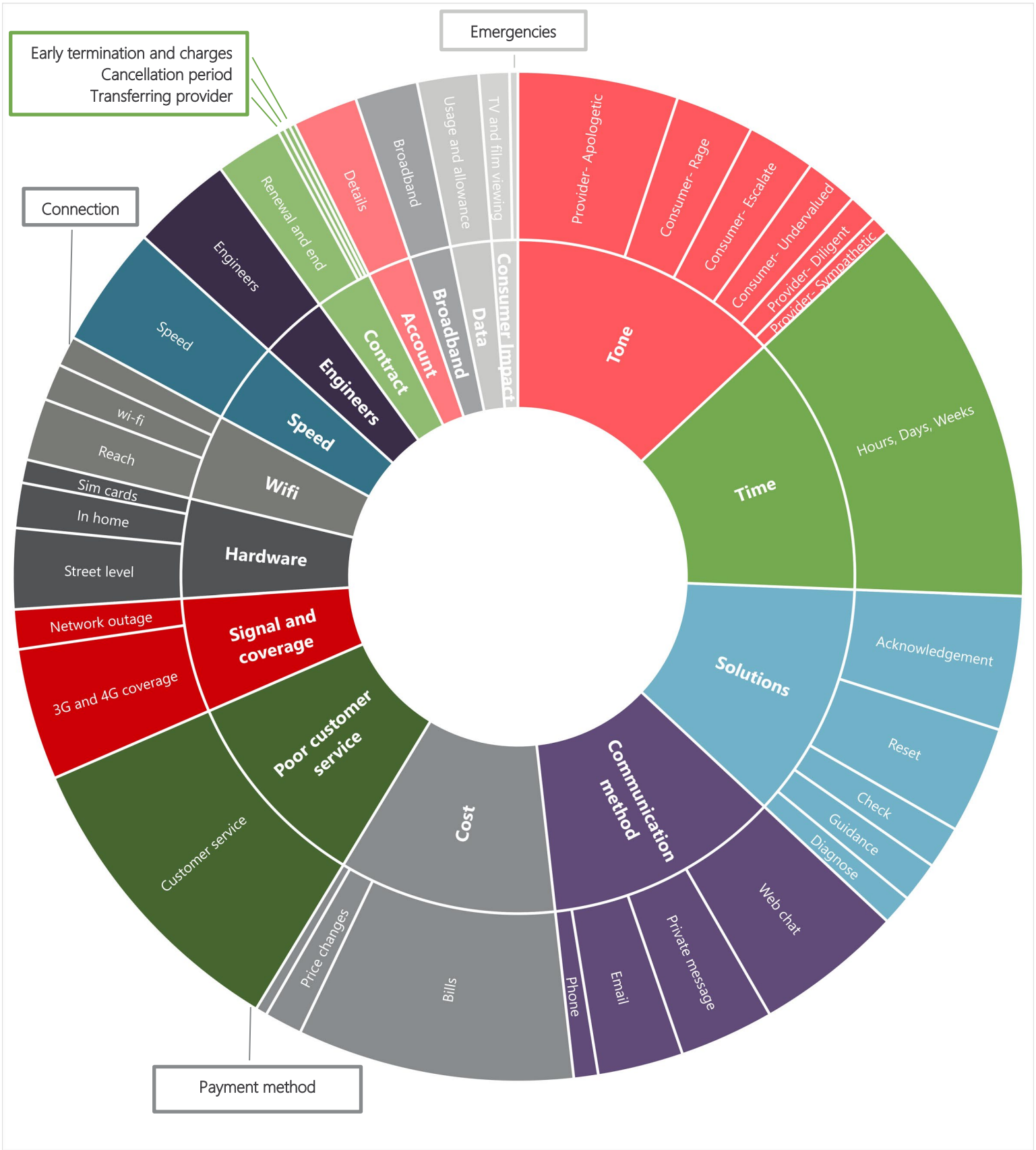
A topic wheel summarises a topic model, and as such is a visual representation of a conversation on social media. In this case, the topic wheel gives an overview of the online conversation around concerns and complaints about fixed and mobile telecoms providers. It is divided into different outer segments (topics), which are made up of smaller inner segments (subtopics). The size of each segment is proportional to the number of posts captured in the corresponding topics and subtopics within the model. The 16 topics and 39 subtopics were created based on the most frequently occurring terms in the data, which were identified using algorithms, then qualitatively reviewed and refined by Ipsos MORI researchers.

From the full cleaned data set, 71% of posts were categorised into the topics described by the topic model and are therefore included in the topic wheel. It is important to note that a single post may appear under multiple topics. The remaining data (29% of posts) could either not be categorised (irrelevant or incoherent), was added to topics which were manually excluded (identified as irrelevant during the qualitative review of the topic model), or formed topics representing less than 0.2% of posts in the overall conversation. Please note that the percentages reported in this chapter are based on all posts in the final clean data set including uncategorised data, unless otherwise stated. It is worth noting that the topic model is generated 'bottom-up' from the data set, rather than being imposed as a thematic framework (as would be the case for other types of qualitative and quantitative analysis).

This means the topics reflect different types of language that are not necessarily equivalent. As such, the topics vary in what they tell us about the social media conversation. For example, some reflect the emotive language included posts (such as the 'Tone' topic), while others relate to specific consumer issues about telecom providers (for example, the 'Signal and coverage' topic). This is a feature of the kind of analysis used to develop the topic model, and is distinct from how a human might group together topics and themes from the data based on a qualitative review (but this would, of course, be a resource-intensive and time-consuming process). The topic model also reflects how few customers explicitly refer to fixed or mobile services within their posts. Although a notable proportion do mention 'Broadband', hence a specific topic for this.

The overall topic model includes retweets, and provider and consumer comments. These are included to give an overview of the online conversation captured by the topic model, as there is value in observing topic size with retweets and provider content included. The next chapter discusses in more detail the topics consumers post about most frequently, including a topic wheel which only contains consumer topics relating to issues with telecoms providers.

Figure 4.1: Overall topic wheel



Please note, the topic wheel reflects the 71% of the final clean data included in the topic model, as outlined in the methodology section, and that a single post may appear under multiple topics.

Table 4.1: Topic and subtopic counts, including retweets, consumer, and provider posts
Note: Rounding results in some discrepancies between subtopic counts and topic totals counts.

Topic	Subtopic	Facebook	Twitter	Total
Time	Hours, Days, Weeks	21%	17%	19%
Tone	Consumer - Rage	5%	3%	4%
	Consumer - Escalate	5%	2%	3%
	Consumer - Undervalued	4%	1%	3%
	Provider - Apologetic	13%	3%	8%
	Provider - Diligent	2%	<1%	1%
	Provider - Sympathetic	2%	<1%	1%
	Tone total	26%	9%	17%
Cost	Bills	18%	9%	13%
	Price changes	2%	2%	2%
	Payment method	1%	<1%	1%
	Cost total	19%	11%	15%
Communication method	Web chat	9%	5%	7%
	Private message	9%	1%	5%
	Email	5%	4%	4%
	Phone	1%	1%	1%
	Communication method total	21%	10%	15%
Poor customer service	Customer service	20%	10%	15%
Solutions	Acknowledgement	9%	4%	6%
	Reset	6%	4%	5%
	Diagnose	1%	2%	1%
	Guidance	3%	1%	2%
	Check	3%	1%	2%
	Solutions total	17%	11%	14%
Signal and coverage	3G and 4G coverage	8%	5%	6%
	Network outage	1%	2%	2%
	Signal and coverage total	9%	7%	8%
Hardware	Street level	3%	4%	4%
	In home	2%	3%	2%
	Sim cards	2%	1%	1%
	Hardware total	7%	8%	7%
Wifi	Reach	3%	3%	3%
	Connection	2%	1%	1%
	Wi-fi	2%	1%	2%
	Wifi total	7%	5%	6%
Speed	Speed	6%	6%	6%
Engineers	Engineers	6%	4%	5%
Contract	Renewal and end	5%	2%	3%
	Early termination and charges	<1%	<1%	<1%
	Cancellation period	<1%	<1%	<1%

	Transferring provider	<1%	<1%	<1%
	Contract total	**6%**	**2%**	**4%**
Data	Usage and allowance	**3%**	**3%**	**3%**
Account	Details	**5%**	**1%**	**3%**
Broadband	Broadband	**3%**	**3%**	**3%**
Consumer impact	TV and film viewing	1%	1%	1%
	Emergencies	1%	<1%	<1%
	Consumer impact total	**2%**	**1%**	**1%**
Uncategorised		**22%**	**36%**	**29%**

The frequencies, even among the largest topics, are relatively small. This demonstrates that no single topic dominates online conversations about fixed and mobile telecoms providers. This also reflects the diversity of social media conversations – there is a huge variety in terms of types of issues and language used.

4.2 'Tone' and 'Time' topics are the backdrop to some of the online conversation

Many posts contained two specific types of language, categorised in the topic model as 'Tone', and references to 'Time'. This highlights the types of terms consumers and providers frequently used in their social media posts. The 'Tone' and 'Time' topics provide the backdrop to some of the specific issues consumers and providers discuss.

'Time' posts specified lengths of time such as minutes, hours, days or weeks. Consumers often spoke about wait times or the length of time they had been affected by particular issues.

"I only got my broadband set up in December and I've been without internet for a week now. We promised an engineer would arrive between 1pm and 6pm today but told now they won't turn up. It's my only day off so I guess I'll have to wait, yet again, for them to visit another time."

Consumer Facebook post

Consumer posts categorised into the 'Tone' topic fall into three subtopics: 'Rage', 'Escalate', and 'Undervalued'. 'Rage' posts tended to be highly emotive, with consumers expressing deep frustration and anger, using terms such as 'utter disgrace' and 'absolute joke'.

"Your service is absolutely disgraceful. No one will answer my calls. A delay is fair enough but waiting 20 minutes is a joke. I am disgusted."

Consumer Twitter post

'Escalate' posts were more likely to include references to official bodies such as 'ombudsman' and 'Ofcom'.

"I made a complaint months ago, still not been resolved. Please can you help me @Ofcom? Annoyed is an understatement!!!"

Consumer Twitter post

'Undervalued' posts contained language which tended to reflect a sense of being let down, with consumers referring to themselves as 'loyal customers'.

"How about lowering your broadband prices for your loyal customers? I pay £50 a month for internet while new customers get it for half that."

Consumer Facebook post

Provider-driven 'Tone' posts were grouped into 'Apologetic', 'Sympathetic' and 'Diligent' subtopics. 'Apologetic' posts were the most common, using terms such as 'sorry' and 'disappointed'.

"Hi [name], I am really sorry to hear you are having issues and have not received the high level of service we pride ourselves on. I apologise for the inconvenience and would like to make this right. Please drop me a private message so I can help further."

Provider Facebook post

'Sympathetic' posts contained terms such as 'understand', 'important' and 'appreciate', alongside adjectives such as 'completely' and 'totally'.

"Sorry to hear of the connection problems you are having [name]. I can understand how frustrating this must be. Please see [weblink] for information."

Provider Twitter post

The language within 'Diligent' posts included reassuring terms like 'assure' and 'work hard' alongside words suggesting immediacy, such as 'soon' and 'quick'.

"Our engineers are working hard to fix the issue and restore the broadband services as soon as possible. We will continue to keep you updated here on Twitter and on our app. Apologies for the inconvenience."

Provider Twitter post

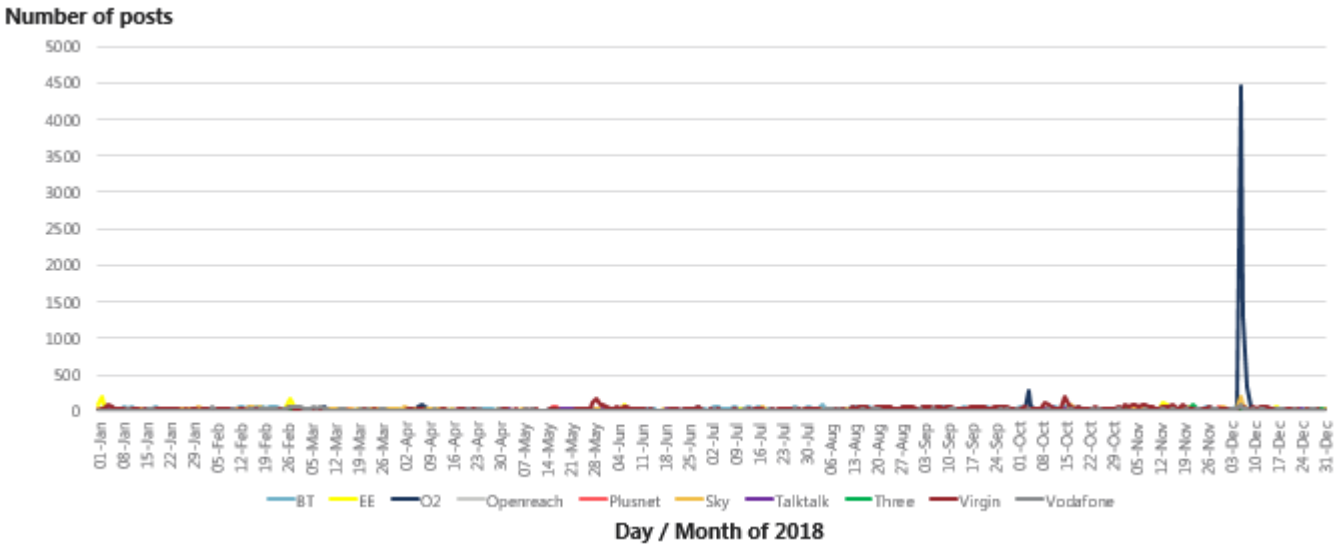
5 Time series analysis

This chapter explores the timing of social media posts made by customers about or to their telecoms providers. Social media activity is notoriously driven by real time events; as such, it is important to consider the extent to which posts made over the course of 2018 were driven by specific telecoms service issues.

Excluding retweets, this research captured a total of 74,379 posts made by customers from January 1st through to December 31st 2018. This translates to an average of 204 per day. Figure 5.1 provides an overview of the number of posts made, by provider, per day.

This figure shows a significant spike of activity on 6/7th December 2018, relating solely to the O2 network outage. Posts referring to O2 on these days account for around ten per cent of the total number of posts across all providers in 2018, and nearly half of all posts about O2.

Figure 5.1 Customer social media posts over time, January – December 2018



Posts by O2 customers relating to the outage were mixed. A common theme, especially in the early hours of 6th, was to ask O2 for clarification of the issue, and to share experiences of whether 4G, texts and calls were working. Most of these comments struck a neutral tone. A further common response was one of frustration and exacerbation at the situation. Above simply being an inconvenience, some customers stressed that the outage was having a significant impact on their business or ability to work.

> "I have over ten phones on O2 for business use. No signal and no data has already causes major issues today. Our management team completely depend on mobile data for a large amount of our workload".
>
> Consumer Twitter post

A request for compensation was also a common theme suggested by some customers. However, this was countered by other customers who criticised their peers for what they felt was a lack of patience and a broader reflection of a societal culture.

> "I hope there's compensation for this inconvenience as I do not want to be charged for 4G when it is not working!"
>
> Consumer Twitter post

"Two things clear on many of these posts – demonstrates perfectly partly what is wrong in modern society, mental meltdown because of no mobile service and an expectation of compensation for it."

Consumer Facebook post

The provider posts from O2 emerged quickly in the day, and asked customers to try switching to 3G, and make the most of wifi. The tone was apologetic and conciliatory, and stressed O2's commitment to fix the problem quickly. Again, customer response to the O2 reaction was mixed. While some felt that O2 should be doing more to proactively contact customers, or to not rely on internet status updates where customers are struggling to get online, others felt that O2 had done a relatively good job at resolving the issue quickly and communicating with customers.

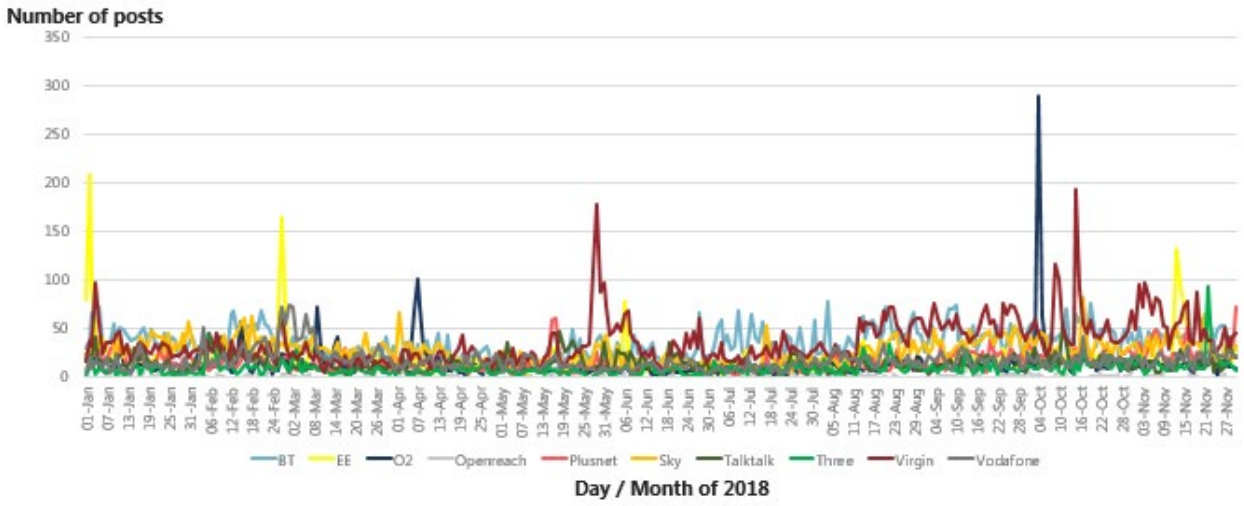
However, the dominance of the outage in early December masks a further series of spikes all driven by telecoms service events. Figure 5.2 shows the total number of customer posts made up to the end of November 2018. In addition to O2, the graph highlights spikes in customer posts for EE, Vodafone, Sky, Virgin Media, BT and Plusnet. These spikes emphasise major short-term problems in delivery of telecoms services, largely driven by service outages.

- Mobile signal outage or widespread problems with broadband connections are the most common cause of spikes in customer posts. These issues explain spikes in the online conversation for EE on 27th February, Sky on 2nd April, Virgin Media on 29th May and 15th October, and O2 on 4th October.
- The largest spike for EE related to downtime in its customer service platforms in early January, such as the chatbot service and telephone lines. This was preventing customers from paying bills, or following up on existing issues that they had previously raised with the customer service team.

However, it should also be noted that in some cases, Facebook threads, rather than specific service events, also drive social media posts made by customers. For example, a conversation on 6th June relating to EE signal/coverage in locations across the UK, and a similar thread on 7th April relating to O2 4G signal and 13th November relating to EE coverage each drive the number of posts for those days.

It is also clear that time-series data does provide some sense of the longevity of service quality issues. For example, posts made about or to Virgin Media become more frequent in August through to November, demonstrating that Virgin Media customers experienced sustained issues relating to connectivity during that period.

Figure 5.2 Customer social media posts over time, January - November 2018



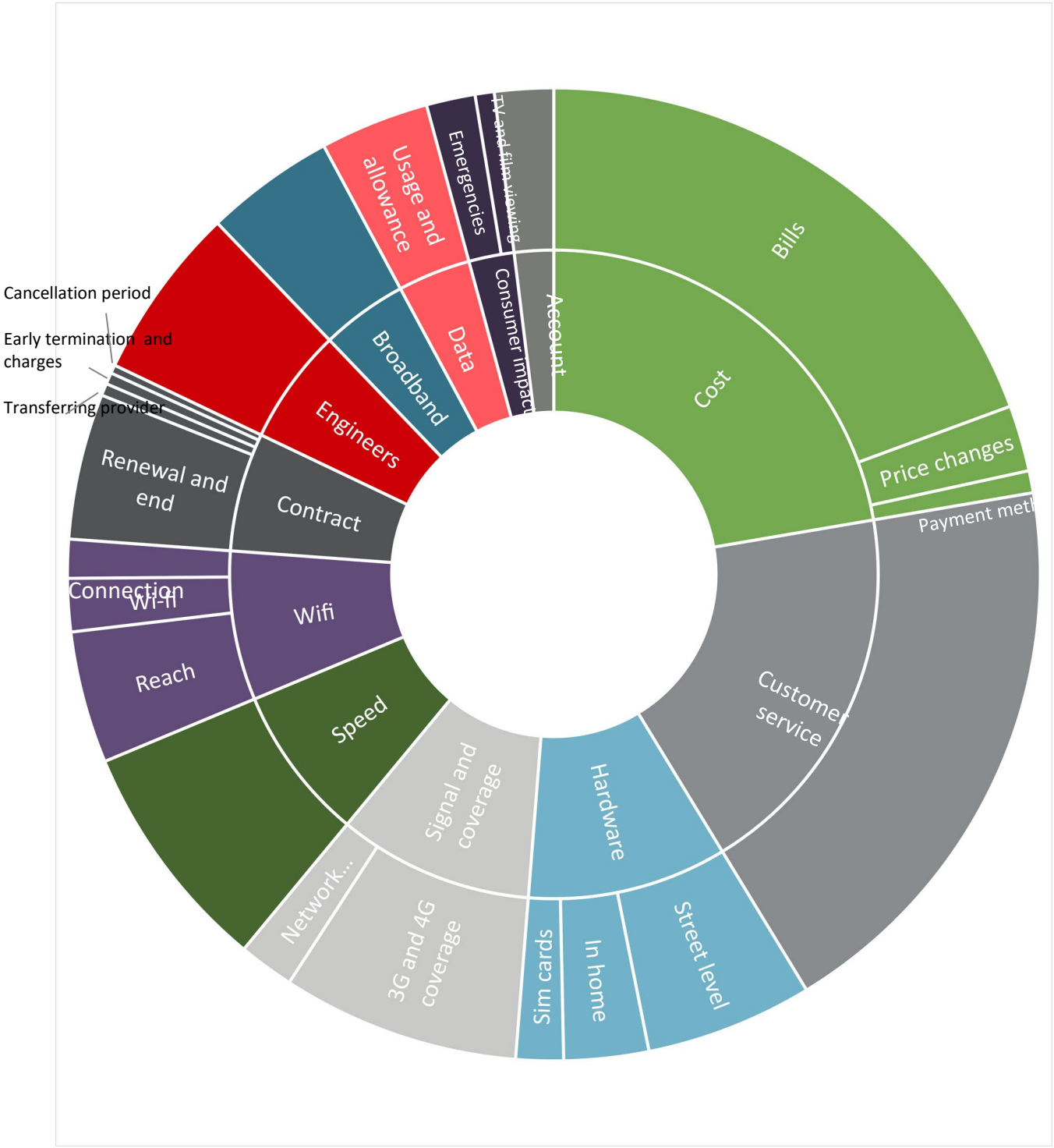
6 Describing the consumer conversation

This chapter describes the concerns and complaints consumers discuss online, highlighting differences between Twitter and Facebook, and how providers respond to consumers.

The following topic wheel displays a subset of topics identified as relating to key consumer issues. The data underpinning this topic wheel contains consumer only posts, excluding retweets, for these topics. This illustrates how some of the key issues are different when looking at consumer only data.

The rest of the chapter explores the range of topics captured by the data, when looking at consumer only posts excluding retweets, but with particular detail on the topics included in the topic wheel.

Figure 6.1: Topic wheel of topics relating to consumer concerns on telecoms providers



Please note that this topic wheel contains a subset of the overall topics in the topic model, selected based on their relevance to consumer issues. The data within each topic reflects consumer only posts, excluding retweets.

6.2 What concerns and complaints do consumers raise on social media?

When analysing consumer-only data (excluding retweets and posts identified as being made by providers) topic frequencies remain broadly similar to the full data set. High levels of 'Tone' and 'Time' posts remain, reflecting the way consumers often frame their concerns and complaints.

Table 6.1: Topics excluding retweets and provider content

Topic	Facebook	Twitter	Total
Time	29%	20%	24%
Cost	25%	12%	17%
Customer service	23%	10%	16%
Tone	20%	8%	13%
Solutions	11%	8%	9%
Signal and coverage	9%	6%	8%
Hardware	8%	9%	8%
Communication method	10%	5%	7%
Wifi	6%	5%	6%
Speed	6%	7%	6%
Engineers	7%	3%	5%
Contract	8%	2%	5%
Data	3%	3%	3%
Broadband	3%	4%	4%
Account	2%	1%	2%
Consumer Impact	2%	2%	2%

In addition to 'Tone' and 'Time', the topics most frequently found in consumer posts were:

- Cost (17%)
- Customer service (16%)
- Solutions (9%)
- Signal and coverage (8%)
- Hardware (8%)

Posts about cost mainly related to 'Bills' (16%), with a much smaller proportion referring to 'Price changes' (2%) and 'Payment methods' (1%). When consumers mentioned bills, they tended to refer to 'monthly' bills and cost terms like 'money', 'pay' and 'charge'. Conversations about price changes related mainly to increases, using terms like 'increase' and 'rise', while payment methods specified 'direct debits', setting up and cancelling payments.

"The network is down so I hope they are going to discount my bill for the inconvenience."

Consumer Twitter post

'Customer service' posts tended to contain descriptive terms such as 'poor' and 'rude' as well as words relating to conversations such as 'speak' and 'tell'.

"It's not just that I can't top up, but no update anywhere or an apology. Very poor customer service you should be ashamed."

Consumer Twitter post

For the 'Solutions' topic (10%), most consumer posts used terms like 'reset', rather than those in the overall topic model that are more likely to be used by providers (e.g. 'guidance' or 'diagnose').

'Signal and coverage' tended to focus on concerns about 3G and 4G. A small proportion (2%) related to network outage, using terms such as 'mobile', 'network' and 'outage'. 'Hardware' posts mentioned sim cards, street level terms (like 'street', 'copper' and 'exchange') and in-home hardware (such as 'plugs', 'sockets' and 'flashing lights' on broadband boxes and hubs).

"The hub was orange for over an hour last night. Went solid green before going back to orange then blue again. Drop outs an internet cut offs like this are really inconvenient"

Consumer Facebook post

Issues relating to 'Speed' (6%), 'Engineers' (5%) and 'Contract issues' (5%) were also raised by consumers. Speed posts contained terms such as 'upload', 'download' and speed measurements like mbps (an abbreviation for megabits per second[16]). Engineer posts contained language about arranging visits, like 'book', 'appointment', and 'come'.

'Contract' issues mainly related to contracts which were being renewed or ended (with terms such as 'end', 'renew' and 'cancel'), in addition to discussions about early terminations and the associated charges ('fee', 'credit', 'charge'). Contract posts also contained terms related to cancelling contracts (with time period terms like '14 days' and '30 days') and transferring providers (including references to requesting PAC codes and transferring mobile phone numbers).

"I am a new customer and went to get a contract for myself but your customer service gave me a business account. It's frustrating. I want to change to a personal contract"

Consumer Facebook post

It is important to acknowledge that the vast majority of posts did not directly refer to fixed or mobile services. This is not unusual for social media data as users often give vague or top level information. This is particularly the case given the data sources used in this research. The Twitter character limit restricts the ability to provide longer detailed posts while Facebook posts were made directly on provider pages, meaning consumers may be expecting providers to already know the context for their issue.

The topic model identified that, while most posts did not explicitly mention fixed or mobile services, a notable proportion did refer to fixed service through mentions of 'Broadband' and 'Wifi'. Across consumer posts, 'Wifi' was more likely to be mentioned (6%) compared to broadband (3%) and specific topics were created for each in the model. 'Wifi' posts contained terms such as 'connection' and 'reach'. 'Broadband' posts included terms like 'fibre' and super optic', alongside speed terms ('fast' and 'slow'). Direct mentions of 'mobile' were more dispersed across other topics (and therefore did not form an independent topic), particularly with in the 'Usage and allowance' and 'Network outage' subtopics, and the 'Phone' and 'Private message' 'Communication method' subtopics.

A small portion of posts detailed the impact of issues with fixed and mobile services on consumers, categorised under the topic 'Consumer impacts'. Of these impacts, consumers tended to focus on 'Emergencies' and 'TV and film viewing'. 'Emergencies' included not being able to contact family members in hospital and school pick up disruptions. 'TV and film viewing' focused on buffering and difficulty streaming videos, particularly on the sites YouTube and Netflix.

[16] https://techterms.com/definition/mbps

"My elderly mum is disabled and has been left with no home phone! It is extremely dangerous for a disabled person not to have a landline! I look forward to hearing from you to resolve this and compensate for the distress caused."

Consumer Facebook post

6.3 Differences between Facebook and Twitter

It is important to recognise that the longer character limit for Facebook posts means they tend to contain richer, more varied, content. In turn, this means posts are more likely to be categorised into multiple topics, resulting in larger proportions of Facebook posts across different topics compared to Twitter. The Facebook data was also captured from public provider pages, suggesting that consumers posting are more likely to want to directly raise issues with their provider (rather than commenting on some aspect of a service in general).

However, while there are some differences in the nature of the consumer comments on Facebook and Twitter, both types of data were analysed as a single dataset to ensure more of the overall conversation on social media about complaints and concerns was captured. The topics raised on Facebook and Twitter were broadly similar, but there were some differences in the issues raised by consumers (looking at consumer only data). 'Customer service', 'Costs', 'Contracts' and 'Engineers', were topics that were all more likely to be raised on Facebook than Twitter.

6.4 How do providers respond to consumers?

Moving away from consumer-only data to look at the full data set excluding retweets, providers often communicate using apologetic language, and look to invite further communication or suggest solutions. Provider posts referred to a range of common 'Communication methods' including 'Webchat', 'Private message', 'Email' and 'Phone'. 'Webchat' and 'Private message' were the most frequently mentioned, suggesting a preference for online communication.

"We are sorry for any inconvenience this has caused and we are glad you have now got this sorted. Please PM us your mobile number and we can raise your feedback internally for further training to avoid this issue in future."

Provider Facebook post

In some instances, providers sought to suggest 'Solutions' to issues in response to consumer complaints. These fell into five categories; 'Acknowledgement', 'Reset', 'Check', 'Diagnose' and 'Guidance'. 'Acknowledgement' posts were the most common, contained terms like 'aware' and 'resolve', 'problem' and 'fault'. 'Reset' posts were the second most common and suggested rebooting systems like Wifi routers.

"We are aware some mobile customers have been experiencing text message issues but this should now be resolved. Can you try turning the handset off and on again?"

Provider Twitter post

'Check', 'Diagnose' and 'Guidance' posts occurred with similar frequencies. 'Check' posts contained terms like 'check', 'status' and 'update'. 'Diagnose' posts contained terms related to identifying issues, such as 'test' and 'run'. 'Guidance' posts often signposted to different information sources, using terms like 'troubleshoot', 'guide' and 'webpagelink' alongside instructing words such as 'follow', 'step' and 'complete'.

"If you are having difficulty with signal there is a link to check your area and see updates on signal here [weblink]. There is also some troubleshooting info that may help you too as changing a few settings can often help resolve this."

Provider Facebook post

6.5 Mentions of Ofcom in conversations

In total, there were 737 references to Ofcom in the data set (including retweets, consumer and provider posts), and 610 in consumer-only data (excluding retweets and provider posts). In the consumer-only data, the distribution of Ofcom mentions broadly reflected the frequency of each topic. For example, Ofcom mentions are most likely to occur in the 'Tone' (29% of the posts that mention Ofcom) and 'Time' (28%) topics. Within 'Tone' posts, Ofcom was most likely to occur within 'Escalating' posts (20%). A relatively small proportion of Ofcom mentions occurred within 'Rage' posts, with just 7% of Ofcom mentions in these posts.

7 How topics group together

This chapter describes in more detail the topics that most often occur together in the online conversation around consumer concerns and complaints about fixed and mobile telecoms.

7.1 Using factor analysis to group topics

Factor analysis groups together related topics from the topic model, based on the likelihood that they occur near each other within individual posts. The findings are based on all the posts included in the topic model, including provider posts and retweets.

Each of the subtopics was grouped into 10 'factors' based on this analysis. These groupings were qualitatively reviewed and refined, creating a series of factor labels. This is a useful way to summarise the topic model and begin to explore the connections between issues in posts by consumers and providers. However, it is worth emphasising that these topics do not always occur together, and that topics from different factors can also occur together.

The % total for each factor reflects the proportion of posts from the total cleaned data set, including retweets and provider content (95,241 posts), which fell into each factor. Posts may be categorised into multiple topics, and therefore factors, meaning totals do not add to 100%. The factors below are ordered from those containing the highest number of posts to the fewest. The topics within each factor are ordered based on topic size, although it is important to note the purpose of the factor analysis is on the relationships rather than the sizes of topics.

Table 7.1: Factor analysis to group topics

Factor	Topic/subtopic
Poor customer service (33%)	Time
	Customer service
	Tone: consumer - Rage
	Tone: consumer - Escalate
	Tone: consumer - Undervalued
	Consumer impacts: Emergencies
Costs and contracts (17%)	Cost: Bills
	Contract: Renewal and end
	Costs: Price changes
	Cost: Payment method
	Contract: Cancellation period
	Contract: Transferring provider
	Contract: Early termination and charges
Sympathetic provider responses - suggesting solutions (14%)	Solutions: Acknowledgement
	Solutions: Check
	Tone: provider – Diligent
Apologetic provider responses - inviting private communication (11%)	Tone: provider – Apologetic
	Communication method: Private message
	Account
	Communication method: Email
	Tone: provider – Sympathetic

Broadband and speed (11%)	Speed
	Hardware: Street level
	Broadband
	Consumer impacts: TV and film viewing
Signal, coverage and data (11%)	Signal and coverage: 3G and 4G
	Data: Usage and allowance
	Hardware: Sim cards
	Signal and coverage: Network outage
Contact details (8%)	Communication method: Webchat
	Communication method: Phone
Wifi (8%)	Solutions: Reset
	Wifi: Connection
	Wifi
Engineers and diagnosing issues (7%)	Engineers
	Solutions: Guidance
	Solutions: Diagnose
Hardware (5%)	Wifi: Reach
	Hardware: In home

7.2 What does the factor analysis tell us?

7.2.1 Poor customer service is emotive

The first factor, 'Poor customer service', demonstrates how issues relating to customer service are particularly emotive. All three of the consumer tones, 'Rage', 'Escalate', and 'Undervalued' were most likely to occur alongside comments about customer service, demonstrating the strength of feeling towards this issue. This factor is the most likely to contain mentions of Ofcom – 56% of the consumer-only posts, excluding retweets, mentioning Ofcom fall within this factor.

Of the 'Tone' subtopics, 'Escalate' (also the subtopic most likely to contain mentions of Ofcom) had the strongest correlation with the topic 'Customer service'. Alongside 'Tone' posts, the factor analysis reflected how consumers often framed their complaints about customer service contextually, by providing details about lengths of 'Time' and the impacts on their lives such as 'Emergencies' (a 'Consumer impacts' subtopic).

> "I am sick of your condescending, rude and unhelpful customer service agents. I understand that faults happen but no service for that length of time is RIDICULOUS."
>
> Consumer Facebook post

7.2.2 Providers use apologetic and sympathetic language, and suggest next steps

The factor analysis resulted in two provider-driven factors. Both reflected providers offering sympathetic responses, coupled with invitations for further private communication or suggestions for immediate solutions to consumer issues. Invitations for further communication were offered in the form of a 'Private message' or 'Email', alongside a request for more details about consumers 'Accounts'. Those suggesting 'Solutions' are most likely to suggest that customer perform 'Checks', while the providers language focuses on acknowledging the issue raised and reassuring the consumer that it will be addressed. This is reflected by the 'Diligent' and 'Acknowledgement' subtopics captured within this factor.

"I'm so sorry that we let you down today. I can understand your frustration, I would feel the same. I assure you once the service is up and running you will receive a full refund for the time you've not had working service. I hope this is sorted for you soon."

Provider Facebook post

Posts containing 'Phone' and 'Webchat' communications channels topics were grouped into one factor, suggesting these are often grouped together.

7.2.3 Costs and contracts often appear together

All cost and contract subtopics were grouped into one factor, reflecting the intertwined nature of these issues. Within this group, the subtopics 'Early termination and charge' and 'Renewal and end' were the most closely related, suggesting that cost and contract issues may be particularly likely to arise at specific times in the customer journey a consumer has with their provider. Within this factor, the 'Price changes' subtopic was most closely related to 'Transferring provider', indicating this as a potential motivation for consumers seeking to change provider.

"So overpriced. Thank god I got my PAC code so I can finally leave after nothing but TERRIBLE service!"

Consumer Facebook post

7.2.4 Broadband and speed are linked to impact on consumers

The broadband and speed factor illustrated how conversations about speed are likely to directly reference broadband. Again, we see consumers speaking about issues in terms of the impact is has on them, illustrated through the grouping of 'Speed' alongside the 'Consumer impact' subtopic 'TV and film viewing'.

"Getting really annoyed now... broadband dropped AGAIN, rubbish picture, haven't been able to watch the football all day because of this. IT NEEDS SORTING TODAY!"

Consumer Facebook Post

"Decent speed and can watch Netflix without buffering. Keep up the good work."

Consumer Twitter Post

7.2.5 Internet and mobile hardware issues are frequently discussed

Direct references to 'Wifi' tend to occur alongside mentions of 'Connection'. Conversations about Wifi are also the most likely to result in in providers suggesting 'Reset' solutions. Mentions about Wifi 'Reach' were likely to occur alongside the 'In home' hardware subtopics, which tends to contain mentions of internet boxes and hubs, and physical connections like sockets and plugs.

'Signal and coverage' topics ('3G and 4G' and 'Network outage') are grouped alongside 'Data' and 'Sim cards' (the 'Hardware' subcategory). 'Signal and coverage' and 'Data' topics were the most closely related topics in this factor, illustrating the interactions between topics predominantly related to mobile service.

"Still having issues with the 4G connection yet they say there's no problems with the new sim cards"

Consumer Twitter post

Mentions of 'Engineers' are often accompanied by provider responses, either suggesting further 'Guidance' or 'Diagnose' solutions. This provides an additional interpretation of the 'Diagnose' subtopic, in that this topic may reflect providers seeking to provide more information about what engineers may do. For example, perform 'tests' or 'diagnose' faults.

"To find out if there is a fault with your service click on this link [weblink] and run a test. If there is a fault you will be able to request an engineer to come out. Let me know how you get on."

Provider Twitter response

7.3 Sentiment analysis

7.3.1 The limitations of sentiment analysis for complex data

A sentiment analysis algorithm was run on the entire data set, including retweets and provider content. It was also applied to consumer only data, excluding retweets. This categorised each post based on whether it contained positive, negative or neutral language. This then created overall sentiment per topic, subtopic and factor. It is important to note that a single post can be categorised into multiple sentiments, which means figures will not sum to 100%.

The length of many of the posts and the wide variety of content and language used means that the value of sentiment analysis for analysing the data is limited. This type of analysis does not always capture the nuances of the English language. For example, the algorithms do not consistently identify sarcasm, or swear words being used to express positive sentiment. Certain thresholds must also be met for posts, or sections of posts, to be categorised under a sentiment meaning not everything will be captured.

7.3.2 Sentiment varies across topics

When analysing the entire data set, including retweets, posts in the majority of topics were, on balance, more negative than positive or neutral. Overall, around a third of posts (32%) were identified as including negative sentiment, and one in five as containing positive sentiments (21%). Almost all posts were found to have some neutral sentiment (94%).

Even so, this analysis does highlight some differences between topics. The topic with the highest proportion of posts containing negative sentiment were 'Customer service' (54% of posts relating to customer service contained negative sentiments). The 'Contract' and 'Bills' topics were also among those with the highest level of negative sentiment (47% and 41% of posts contained negative sentiments respectively). The 'Transferring providers' subtopic of the 'Contract' topic was particularly negative, with 51% of posts in this subtopic containing negative sentiment. Unsurprisingly, negative sentiment was also high amongst the 'Consumer impacts' topic (44%) which contained posts relating to how disruptions or poor service had impacted consumers lives. Of the consumer specific 'Tone' subtopics, negative sentiment was highest among 'Escalate' posts, with 60% of posts in this subtopic containing negative sentiment.

There were a small number of topics where the level of positive sentiment was higher than negative sentiment – this was predominantly within provider specific topics. For example, 32% of post within the 'Diagnose' subtopic of 'Solutions' were identified as containing positive sentiment (compared to 25% of posts within this subtopic containing negative sentiment). For the 'Phone' subtopic of 'Communication methods', 31% of posts were identified as containing positive sentiment, compared to 24% of posts containing negative sentiments. There were also more posts containing positive sentient compared to negative sentiment in the 'Early termination and charges' subtopic of the 'Contract' topic (39% positive vs 23% negative).
A similar picture emerged for sentiment analysis of consumer only posts, excluding retweets. For these posts, 32% were identified as containing negative sentiment, 17% of as containing positive sentiment and 93% as containing neutral sentiments. 'Customer service' remained the topic with the highest level of negative sentiment (59%), followed by

'Contract' (47% of post contained negative sentiments). Again, of the 'Contract' subtopics, 'Transferring providers' was the most negative, with 57% of post containing negative sentiments compared to 14% of posts containing positive sentiments. The 'Early termination and charges' subtopic of 'Contract' remained more positive (38% of posts contained positive sentiment) than negative (25%).

Sentiment analysis was also applied to the factors (which related to all posts in the clean dataset, including retweets, provider and consumer posts). When looking at sentiment by factor, rather than by topic, all factors contain more negative sentiments than positive. The highest level of negative sentiment was within the 'Sympathetic provider responses – suggesting solutions' factor (48% of posts in this factor were identified as containing negative sentiments compared to 27% containing positive sentiments). There are different ways this could be interpreted - for example, providers may respond to consumers using language identified as more negative as they suggest as they suggest immediate troubleshooting steps. 'Poor customer service' contained the second highest level of negative sentiment (45%) followed by 'Cost and Contracts' (40%).

8 Provider differences

This chapter summarises the overall factor analysis by provider then discusses the largest topics and relevant issues for each of the nine providers and Openreach.

8.1 Share of posts by platform

The table below provides an overview of the share of posts, by platform, for each provider.

Table 8.1: Total posts, excluding retweets by provider

		BT	EE	O2	Openreach	Plusnet	Sky	TalkTalk	Three	Virgin Media	Vodafone
Data source	Facebook	39%	58%	68%	53%	90%	18%	66%	76%	12%	71%
	Twitter	61%	42%	32%	47%	10%	82%	34%	24%	88%	29%

8.2 Interpreting provider differences

There are a number of interlinked reasons that could help explain some of the differences in findings between different providers, including the number of posts for different topics and how the factors vary across providers. These include:

- Provider size and range of services: The providers included in this research vary greatly in their size and market share, and also offer different services. For example, some only provide fixed services, while others have both mobile and fixed services and offer a wide range of products within each.
- Customer demographics: Different providers have different types of customers. Certain demographics vary, as will their propensity to be online and use social media.
- Provider online presence: There are differences between the online presences different providers have. This includes variations in the number of Twitter followers and Facebook users visiting their Facebook pages, as well as differences in the extent to which providers encourage and interact with consumers on these platforms.
- Data capture method: While the search queries used to capture data from Twitter and Facebook were designed to be as consistent and relevant as possible, there are limitations to this. For example, some provider names required more exclusions to be added to the search query.

For these reasons, it is not appropriate to simply make comparisons between providers when reviewing findings. For example, it cannot be inferred that the provider with the largest number of posts on a particular topic is necessarily causing more concerns and complaints among its overall customer base. As such, the analysis by provider explores how the factors vary across providers and the key topics captured for each provider, rather than making simple comparisons between providers.

8.3 Key consumer factors by telecoms providers

In this section, we outline how frequently the 10 overall factors appear in a subset of the data – based on consumer posts, excluding retweets and provider content. This allows us to begin to build a picture of the consumer conversation about telecoms complaints and concerns online, based on focusing on the seven factors that relate most directly to consumer experience. Table 8.3 sets out these seven consumer-driven factors by provider, ranked by how common each factor is overall.

Table 8.2: Overall factors based on consumer data excluding retweets, by provider

Factor	BT	EE	O2	Openreach	Plusnet	Sky	TalkTalk	Three	Virgin Media	Vodafone	Total
Customer service	40%	33%	30%	30%	**50%**	34%	46%	34%	37%	47%	38%
Costs and contracts	20%	21%	18%	10%	**26%**	18%	23%	21%	15%	24%	20%
Broadband and speed	20%	11%	2%	46%	16%	**19%**	18%	7%	8%	11%	13%
Signal, coverage and data	4%	25%	25%	6%	4%	5%	4%	**30%**	5%	16%	11%
Wifi	7%	7%	6%	3%	**9%**	5%	8%	7%	7%	6%	7%
Hardware	10%	4%	1%	**14%**	6%	7%	8%	3%	6%	3%	6%
Engineers and diagnosing issues	9%	2%	1%	9%	8%	6%	**10%**	1%	5%	3%	5%

Care should be taken when drawing simple comparisons across providers, for the reasons outlined earlier in this section. These findings summarise the online conversation related to each provider, rather than being a robust basis for comparing their performance on key issues. However, the factors allow us to see where it is relatively more or less common for consumers to discuss certain issues about different providers online.

- The findings highlight that 'Poor customer service' was the most common factor for all providers. Customers often used social media to comment on the quality of the service they have received and to attempt to resolve issues. This includes posts discussing the time taken to fix issues, and expressing dissatisfaction with their experience as they interact with customer service teams. The online conversation about Plusnet was more focused on this factor relative to other issues.
- Similarly, 'Costs and contracts' issues were raised by consumers across providers. The online conversations about Plusnet, TalkTalk and Vodafone contained relatively more posts about these issues.
- Some factors are service specific, related either to mobile or fixed services. These include the 'Broadband and speed' factor, and the 'Signal, coverage and data' factor which were relatively important for fixed and mobile providers respectively. This emphasises the importance of ensuring even broad comparisons are made only between providers of similar services.
- For the 'Signal, coverage and data' factor, this may have the potential to be refined to reflect actual performance among mobile providers. Based on the time series data, O2 and EE both had significant outages during 2018, with big spikes of complaints and concerns, and this factor is more common in the online conversation about these providers. This is a smaller issue in the online conversation about Vodafone. However, it is interesting to note that Three had the highest proportion of posts relating to 'Signal, coverage and data', despite not having a major outage, reflecting how this issue was key to consumers regardless of outage events.
- Similarly, the 'Hardware' and 'Engineers and diagnosing issues' factors were more common in the online conversation about providers of fixed services. The latter factor was more common for TalkTalk, Openreach and BT.
- In the case of the 'Broadband and speed' factor, posts about Openreach were largely driven consumers asking or complaining about superfast broadband not being available in their area.

8.4 Key topics for telecoms providers

A topic summary for each provider is described below. First, the text describes the number of posts for each provider captured in the full data set, *including* retweets, provider and consumer data. This is also given as a percentage of all posts contained within the full data set, to demonstrate how the overall data set breaks down between the providers.

The most frequent topics for each provider are then set out in separate tables. However, these tables are based on the clean data set excluding retweets (but including both provider and consumer posts). Retweets are excluded from provider-level analysis for the reasons outlined in the 'Methodology' chapter. Each table contains the 10 topics with the highest number of posts for each of the nine telecoms provider. Specific analysis of Openreach is included at the end of this section.

For each table the 'total' column reflects the proportion of posts captured within the relevant topic when looking at the entire clean dataset, excluding retweets and including provider and consumer posts. The provider specific column (e.g. 'BT') includes the number of posts captured by the relevant topic when looking at the posts specific to that provider (excluding retweets, but including provider and consumer posts).

When looking across the key topics by provider, most have more posts captured from Facebook rather than Twitter. However, Sky and Virgin Media posts were predominantly captured through Twitter, suggesting a preference for these providers or their consumers to communicate through Twitter. As discussed in the previous section, the different balance of posts across different social media makes drawing simple comparisons between providers challenging. Instead, they are included here to illustrate the relative importance of different topics for each provider separately.

Overall summary

Across all providers, 'Customer service' was the most common consumer specific topic, when looking at the top three consumer specific topics per provider. The most frequently mentioned topics for a specific provider were often closely aligned to the providers' services. For example, 'Signal and coverage' being more common for providers who focus more on mobile services.

There were differences in the extent to which providers communicated with their customers publicly on Facebook and Twitter, indicated by the 'Communication methods' and 'Solutions' provider specific topics. EE, Sky, Virgin Media and Vodafone all had these topic in their top three (excluding 'Tone' and 'Time' topics which are more general), suggesting they may be more likely to respond to consumers on social media than other providers.

BT

Posts relating to BT were more common on Twitter (61%) than on Facebook (39%).

Key consumer topics for BT included 'Cost', 'Hardware' and 'Customer service'. The inclusion of 'Hardware', which references 'fibre optics' and 'copper cables' and 'street cabinets', is likely to reflect the fixed services BT provides. It is possible that this may also reflect consumers confusing BT with Openreach, as Openreach is a division of BT which owns and is responsible for fibres, wires and cables used to connect broadband and telephone networks.

8.4.1 Top ten consumer topics for BT

Topic	BT	Total
Time	23%	19%
Costs	16%	15%
Tone	15%	18%
Hardware	14%	7%
Customer service	13%	15%
Communication methods	13%	15%
Solutions	10%	14%
Speed	9%	6%
Wifi	8%	6%
Engineers	8%	5%

EE

These conversations were slightly more likely to happen on Facebook (58%) than on Twitter (42%).

Two of the three largest topics for EE were provider specific topics ('Communication methods' and 'Solutions'), suggesting EE may frequently interact with their customers through social media and directly reply to complaints or concerns raised.

8.4.2 Top ten consumer topics for EE

Topic	EE	Total
Communication methods	22%	15%
Customer service	19%	15%
Solutions	17%	14%
Tone	17%	18%
Signal and coverage	16%	7%
Costs	14%	15%
Time	12%	19%
Speed	6%	6%
Account	6%	3%
Hardware	5%	7%

O2

These conversations were more likely to happen on Facebook (68%) rather than Twitter (32%).

'Signal and coverage' was the largest customer topic for O2, followed by 'Cost' and 'Customer service'. 'Signal and coverage' was also the topic containing the largest number of posts overall, reflecting O2's provider type as that of mobile services, but this predominantly came as a result of the network outage O2 suffered in December 2018 Posts referring to O2 on the 6th and 7th of December 2018 when the outage occurred, accounted for around 10% of the total number of posts across all providers in 2018. The outage impacted customer's ability to use the network, which included not being able to get signal on their mobile phones.

8.4.3 *Top ten consumer topics for O2*

Topic	O2	Total
Signal and coverage	18%	7%
Tone	16%	18%
Costs	14%	15%
Time	13%	19%
Customer service	13%	15%
Solutions	12%	14%
Communication methods	9%	15%
Data	5%	3%
Contract	4%	4%
Wifi	4%	6%

Plusnet

Plusnet posts were more likely to occur on Facebook (90%) compared to Twitter (10%).

'Solutions' and 'Communication methods' were the largest topics for Plusnet. This suggested high levels of interactions with customers through Facebook and Twitter as both topics are predominantly based on provider posts. 'Cost' and 'Customer service' were the largest consumer topics, in line with the topic consumer issues across all providers.

8.4.4 Top ten consumer topics for Plusnet

Topic	Plusnet	Total
Time	29%	19%
Tone	27%	18%
Communication methods	20%	15%
Costs	20%	15%
Solutions	19%	14%
Customer service	19%	15%
Hardware	9%	7%
Wifi	9%	6%
Speed	9%	6%
Engineers	6%	5%

Sky

Posts about Sky were more likely to occur on Twitter (82%) than Facebook (18%).

'Solutions' and 'Communication methods' were among the most common topics, reflecting Sky's interaction with customers on Facebook and Twitter. 'Cost' and 'Customer service' were the topic consumer topics, followed by 'Engineers' and 'Hardware'.

Table 8.3: Top ten consumer topics for Sky

Topic	Sky	Total
Solutions	16%	14%
Time	16%	19%
Communication methods	15%	15%
Tone	14%	18%
Costs	13%	15%
Customer service	12%	15%
Engineers	10%	5%
Hardware	9%	7%
Speed	7%	6%
Wifi	5%	6%

Talktalk

There were almost twice as many posts captured in the Facebook data (66%) compared to Twitter (34%).

'Cost' and 'Customer service' were the largest topics, demonstrating customer's key concerns. This was followed by similar proportions of posts falling into provider topics - 'Communication methods' and 'Solutions'.

Table 8.4: Top ten consumer topics for Talktalk

Topic	Talktalk	Total
Time	25%	19%
Tone	17%	18%
Costs	16%	15%
Customer service	16%	15%
Communication methods	16%	15%
Solutions	11%	14%
Speed	8%	6%
Hardware	8%	7%
Engineers	8%	5%
Wifi	7%	6%

Three

More Three posts were captured on Facebook (76%) than Twitter (24%).

Of the two topics containing the largest number of posts, one related to providers and one to consumers. The largest topic was 'Solutions', suggesting Three often make posts suggesting immediate solutions to issues raised by consumers. The second largest topics was 'Signal and coverage', reflecting Three's mobile services offer.

Table 8.5: Top ten consumer topics for Three

Topic	Three	Total
Signal and coverage	22%	7%
Solutions	21%	14%
Tone	21%	18%
Customer service	16%	15%
Costs	14%	15%
Time	13%	19%
Communication methods	10%	15%
Data	7%	3%
Contract	6%	4%
Wifi	6%	6%

Virgin Media

Virgin Media posts were more likely to be taken from Twitter (88%) than Facebook (12%), suggesting a lower level of engagement between Virgin Media and consumers on Facebook.

'Customer service' and 'Cost' were the largest topics, following the broader framing topic of 'Time'. This is in line with the consumer issues raised across all providers.

Table 8.6: Top ten consumer topics for Virgin Media

Topic	Virgin Media	Total
Time	24%	19%
Customer service	14%	15%
Costs	13%	15%
Solutions	12%	14%
Tone	11%	18%
Wifi	6%	6%
Hardware	6%	7%
Communication methods	5%	15%
Engineers	4%	5%
Speed	4%	6%

Vodafone

Conversation related to Vodafone were more common on Facebook (71%) than Twitter (29%).

The largest topic for Vodafone was 'Communication methods'. This, combined with 'Solutions' the largest topic for Vodafone, may reflect how Vodafone interacts with consumers on Facebook and Twitter. 'Customer service' is the largest consumer topic, followed by 'Cost' and 'Account'.

Table 8.7: Top ten consumer topics for Vodafone

Topic	Vodafone	Total
Communication methods	32%	15%
Tone	32%	18%
Customer service	19%	15%
Solutions	18%	14%
Time	17%	19%
Costs	15%	15%
Account	9%	3%
Signal and coverage	8%	7%
Speed	6%	6%
Contract	6%	4%

8.5 Key topics for Openreach

As noted, Openreach differs from the telecoms providers analysed within this research. Openreach is a subsidiary of British Telecommunications plc with a network role related hardware (they are responsible for items such as cables and wires used to connect telephone and broadband services). While they don't have end consumers themselves, as providers are their customers, end consumers may have contact with Openreach. As interactions with Openreach may form part of a consumers experience of their telecoms provider it has been included, although it is important to view Openreach separately from the other providers given these reasons.

Conversations relating to Openreach were relatively evenly split between Facebook (53%) and Twitter (47%). 'Hardware', 'Broadband', 'Speed' and 'Wifi' were the top consumer issues for Openreach, reflecting their role in maintaining the UK's broadband infrastructure.

Topic	Openreach	Total
Hardware	34%	7%
Time	17%	19%
Broadband	15%	3%
Speed	12%	6%
Wifi	12%	6%
Solutions	10%	14%
Tone	9%	18%
Engineers	9%	5%
Costs	8%	15%
Customer service	8%	15%

9 Appendices

Queries

Provider	Full query
BT	((("british telecom*" OR BT OR "@bt" OR "@bt_uk" OR "@btcare" OR "@btgroup" OR "@btbusinesscare") NEAR/1 (Wifi OR "wi-fi" OR mobile OR network OR provider OR service OR infinity OR homephone OR "home phone" OR landline OR "land line" OR broadband OR "broad band" OR internet OR hub OR "home hub" OR homehub OR fiber OR fibre OR 3g OR 4g)) OR (btinfinity OR bthomephone OR btmobile OR btbroadband OR btinternet OR bthub OR btvision OR "bt infinity" OR btinfinity OR "bt superfast")) NOT ("@btsports" OR "@BTscotland" OR "btsportfootball" OR "@btsportscore" OR "@btsportboxing" OR "@btsportufc" OR "@BTsportbars" OR "bt sport" OR "bt sports" OR BTsport OR order OR sale OR sales OR deal OR deals OR promo OR giveaway OR "give aways" OR save OR offer OR offers OR voucher OR news OR ad OR advert OR unlock OR dogs OR dog OR "internet of things" OR IoT OR win)
TalkTalk	((("talk talk" OR talktalk OR "@talktalk" OR "@talktalktv" OR "@talktalkgroup" OR "talktalkgroup" OR "@talktalkbusiness" OR talktalkbusiness OR "@TTB_partner" OR "ttb_partner" OR "@talktalkbizcare" OR talktalkbizcare OR talktalkmobile OR "@talktalkufoyork" OR "@talktalkbiz") NEAR/1 (Wifi OR "wi-fi" OR mobile OR network OR provider OR service OR homephone OR "home phone" OR landline OR "land line" OR broadband OR "broad band" OR internet OR hub OR "home hub" OR homehub OR fiber OR fibre OR 3g OR 4g)) OR ("talktalk unlimited")) NOT ("talk talk talk" OR talktalktalk OR order OR sale OR sales OR deal OR deals OR promo OR giveaway OR "give aways" OR save OR offer OR offers OR voucher OR news OR ad OR advert OR unlock OR dogs OR dog OR "internet of things" OR IoT OR win)
Sky	((bskyb OR skyuk OR "sky uk" OR skyhelpteam OR sky OR "sky talk" OR skytalk OR "@skyuk" OR "@skyhelpteam" OR skymobile OR "@sky") NEAR/1 (Wifi OR "wi-fi" OR mobile OR network OR provider OR service OR homephone OR "home phone" OR landline OR "land line" OR broadband OR "broad band" OR internet OR hub OR "home hub" OR homehub OR fiber OR fibre OR 3g OR 4g) OR ("sky max")) NOT ("night sky" OR skies OR "blue sky" OR "sky news" OR "sky sports" OR "sky sport" OR "sky-high" OR skyhigh OR "sky high" OR skynews OR "sky Q" OR skysport OR "@skysports" OR "@skynews" OR "@skynewsbreak" OR "@skysportsnews" OR "@skysportsstatto" OR "@skynewsbiz" OR "@skynewspolitics" OR "@skynewsboxing" OR "@skyfootball" OR "Sky atlantic" OR "sky one" OR "sky movies" OR order OR sale OR sales OR deal OR deals OR promo OR giveaway OR "give aways" OR save OR offer OR offers OR voucher OR news OR ad OR advert OR unlock OR dogs OR dog OR "internet of things" OR IoT OR win)

Virgin Media	(((virgin OR "virgin media" OR virginmedia OR "virgin cable" OR "virgin media cable" OR "@virginmedia" OR "@virginmediacorp" OR "@vmbusiness" OR virginmobile) NEAR/1 (Wifi OR "wi-fi" OR mobile OR network OR provider OR service OR homephone OR "home phone" OR landline OR "land line" OR broadband OR "broad band" OR internet OR hub OR "home hub" OR homehub OR fiber OR fibre)) OR ("virgin superhub" OR "virgin super hub")) NOT (virgintrains OR virgingalactic OR "virgin mary" OR naked OR "virgin atlantic" OR virginatlantic OR train OR sex OR married OR "Virgin Atlantic" OR holidays OR travel OR "Virgin Active" OR "Virgin Racing" OR "@virginmoney" OR order OR sale OR sales OR deal OR deals OR promo OR giveaway OR "give aways" OR save OR offer OR offers OR voucher OR news OR ad OR advert OR unlock OR dogs OR dog OR "internet of things" OR IoT OR win)
EE	(((ee OR everythingeverywhere OR "everything everywhere" OR "@ee") NEAR/1 (Wifi OR "wi-fi" OR mobile OR network OR provider OR service OR homephone OR "home phone" OR landline OR "land line" OR broadband OR "broad band" OR internet OR hub OR "home hub" OR homehub OR fiber OR fibre OR 3g OR 4g)) OR ("4gee home" OR "4g ee home")) NOT (unlocked OR job OR jobs OR sprint OR win OR selling OR order OR sale OR sales OR deal OR deals OR promo OR giveaway OR "give aways" OR save OR offer OR offers OR voucher OR news OR EEgoldensim OR ad OR advert OR unlock OR dogs OR dog OR "internet of things" OR IoT OR win)
O2	((o2 OR 02 OR telefonica OR telephonica OR telefonica OR "@o2" OR "@02" OR "@o2businessuk") NEAR/1 (Wifi OR "wi-fi" OR mobile OR network OR provider OR service OR homephone OR "home phone" OR landline OR "land line" OR broadband OR "broad band" OR internet OR hub OR "home hub" OR homehub OR fiber OR fibre OR 3g OR 4g)) NOT ("@theo2" OR "@the02" OR "the o2" OR "the 02" OR "presale" OR "o2 academy" OR "02 academy" OR "@o2academybrix" OR "@o2sbe" OR "@o2islington" OR "@o2forumktown" OR "@O2academyoxford" OR "@o2music" OR order OR sale OR sales OR deal OR deals OR promo OR giveaway OR "give aways" OR save OR offer OR offers OR voucher OR news OR ad OR advert OR unlock OR dogs OR dog OR "internet of things" OR IoT OR 02/ OR /02/ OR 02: OR win)
Three	((("@threeuk" OR "@3" OR "@three" OR "@threeuksupport" OR threeuk OR "three uk" OR Hutchinson OR "Hutchinson 3g" OR H3g OR "three mobile" OR 3mobile OR "3 mobile" OR 3uk OR "3 uk" OR "@threeuk" OR "@3" OR "@three" OR "@threeuksupport") NEAR/1 (Wifi OR "wi-fi" OR mobile OR network OR provider OR service OR homephone OR "home phone" OR landline OR "land line" OR broadband OR "broad band" OR internet OR hub OR "home hub" OR homehub OR fiber OR fibre OR 3g OR 4g)) OR ("three homefi" OR "three home fi" OR mifi OR "mi-fi" OR "mi fi" OR "mi-fi")) NOT ("level 3" OR "@virginmedia" OR "virgin media" OR virginmedia OR times OR fraudster OR order OR sale OR sales OR deal OR deals OR promo OR giveaway OR "give aways" OR save OR offer OR offers OR voucher OR news OR ad OR advert OR unlock OR dogs OR dog OR "internet of things" OR IoT OR win)

Vodafone	(((vodaphone OR vodaphoneuk OR vodafone OR "voda-phone" OR "voda-fone" OR "vodaphonegroup" OR "@vodafoneuk" OR "@vodafonegroup" OR "@vodaphoneukcomms" OR "@vodafoneukbiz" OR vodafonegroup OR vodafoneuk) NEAR/1 (Wifi OR "wi-fi" OR mobile OR network OR provider OR service OR homephone OR "home phone" OR landline OR "land line" OR broadband OR "broad band" OR internet OR hub OR "home hub" OR homehub OR fiber OR fibre OR 3g OR 4g)) OR ("vodaphone superfast" OR "vodafone superfast")) NOT (vodafonetechstarter OR "Vodaphone tech starter" OR ad OR advert OR "v-home" OR "v home" OR order OR sale OR sales OR deal OR deals OR promo OR giveaway OR "give aways" OR save OR offer OR offers OR voucher OR news OR unlock OR dogs OR dog OR "internet of things" OR IoT OR win)
Plusnet	(((plusnet OR "plus net" OR "plus-net" OR "@plusnet" OR "@plusnethelp" OR plusnethelp) NEAR/1 (Wifi OR "wi-fi" OR mobile OR network OR provider OR service OR homephone OR "home phone" OR landline OR "land line" OR broadband OR "broad band" OR internet OR hub OR "home hub" OR homehub OR fiber OR fibre OR 3g OR 4g)) OR ("plusnet unlimited" OR "plusnet unlimited")) NOT (order OR sale OR sales OR deal OR deals OR promo OR giveaway OR "give aways" OR save OR offer OR offers OR voucher OR news OR ad OR advert OR unlock OR dogs OR dog OR "internet of things" OR IoT OR win)
Openreach	((Openreach OR "open reach" OR "bt openreach" OR "bt open reach" OR btopenreach) NEAR/1 (Wifi OR "wi-fi" OR mobile OR network OR provider OR service OR homephone OR "home phone" OR landline OR "land line" OR broadband OR "broad band" OR internet OR hub OR "home hub" OR homehub OR fiber OR fibre OR 3g OR 4g)) NOT (order OR sale OR sales OR deal OR deals OR promo OR giveaway OR "give aways" OR save OR offer OR offers OR voucher OR news OR ad OR advert OR unlock OR dogs OR dog OR "internet of things" OR IoT OR win)

Daniel Cameron
Research Director
daniel.cameron@ipsos.com

Steven Ginnis
Research Director
steven.ginnis@ipsos.com

Olivia Lohoar Self
Senior Research Executive
olivia.lohoarself@ipsos.com

Lore Bizgan
Research Executive
lore.bizgan@ipsos.com

For more information

3 Thomas More Square
London
E1W 1YW

t: +44 (0)20 3059 5000

www.ipsos-mori.com
http://twitter.com/IpsosMORI