EXPLORING THE VALUE OF MOBILE PHONE ACTIVITY HISTORY IN IDENTITY ASSURANCE

ALPHA PROJECT OUTPUT WHITE PAPER

ANDREW TOBIN
JULY 2016

Executive Summary

This document details the results of the OIX Activity History Alpha Project.

This project was created to examine the usefulness of new sources of activity data in order to assist the Verify identity providers' ("IDPs") approval process for registering new GOV.UK Verify (henceforth "Verify") users. Activity history is an important component of the Verify identity proofing process. Mobile phone activity history has been identified as a particularly useful attribute.

As no activity data is available directly from mobile operators in the UK yet, TeleSign worked with the Cabinet Office's Governmetn Digital Service (GDS), Digidentity and the Post Office to use TeleSign data as a proxy for mobile operator data. TeleSign's data originates from security messages sent to and from UK mobile phones by Telesign on behalf of their global customers including 20 of the world's top 25 web properties.

The teams initially did a batch test to determine the percentage of UK mobile phone users who have interacted with TeleSign's services, and then built a real-time interface to obtain activity history from TeleSign. This data was then used to determine how many new Verify registrants who failed to get an LOA2 account could have succeeded with this additional activity data. Additionally, a projection was made using this information to determine what the impact would be were mobile operators to provide activity history information.

The phone numbers of approximately 9000 new Verify registrants who had no other activity information were sent to TeleSign in real time from Digidentity and Post Office. An average of 33% of these numbers had used TeleSign's services and therefore had activity recorded within a 90 day period. For the purposes of this project, it was agreed that TeleSign data was ranked "medium" quality with a "short" tenure period (described in detail in the report). When used to calculate the virtual increment to the user's approval score (the user's actual score was not changed in this project), approximately 10% of those users were within a small margin of achieving approval and would therefore have been uplifted from a fail to a pass. This equates to an additional approximately 300 users who would have been approved for a Verify LOA2 account.

When projected forwards to the potential for activity history data direct from UK mobile operators, it was determined that 93% of users could get increased scores if mobile operator data was used, compared to the 33% from TeleSign data. Additionally, it was estimated that 57% would get the highest possible activity history score due to the length of time they had been with their mobile operator.

When viewed from an end user perspective, this means that, of the approximately 9000 users who failed to get a Verify account, 3000 could have their scores increased using TeleSign data and 8200 could have their scores increased (and by

a significantly higher margin) using direct mobile operator data. Using this information, IDPs will be able to determine the actual uplift from fail to pass within their systems using their own proprietary scoring methodologies.

Background

UK government-approved Identity Providers (IDPs) need to confirm if registrants are real, live people who can be proven to be who they say they are. In doing so, the government issued standard GPG 45 requires that the IDPs check information in 3 categories: Citizen (such as national insurance details), Money (such as credit reference data) and Living (such as mobile phone information). In the latter category, providing recent information to confirm the asserted identity belongs to someone who is active and behaving like a "normal" human being is an important factor. This is called "Activity History".

Sources of suitable activity history have proven somewhat hard to come by. This project was created in order to investigate additional sources, and stimulate those sources that do have good data. If, by utilizing new sources of data, the approval rate for new Verify applicants can be increased, it will be a positive result. If this in turn stimulates known data originators to open up access to their data to help their customers to gain approved Verify accounts, this will also be viewed positively.

Hypothesis

Identity Providers could meet the requirements for GPG 45 Element E (Activity History) by:
- linking a user's asserted identity details to a mobile telephone number through a robust process;
- then calling a bespoke service provided by TeleSign that leverages the user's activity associated with the mobile phone number to demonstrate 'normal' activity history.

The combination of services would result in an increase in the GOV.UK Verify success rate and reduce the time users need to register a digital identity with an Identity Provider.

Context

TeleSign provides services to protect online user security to 9 of the 10 largest websites in the US and 20 of the 25 largest websites globally. TeleSign enables its clients to bind a mobile phone number to the end user's account by verifying the user is in possession of the phone, using multiple 2-factor authentication mechanisms. TeleSign also provides additional information about mobile phone numbers that can influence the risk level assigned to a mobile phone number.

Example use case: When a user logs in to a website of a TeleSign client, or changes their password or carries out other security sensitive activities, a request is made to the TeleSign service to calculate a risk score associated with information

pertaining to the mobile phone number. The website takes proportionate action based upon the risks it perceives which may include sending an SMS message or other mechanisms for performing two factor authentication. Note that TeleSign only sees activity data that passes through TeleSign systems as a result of TeleSign client requests such as in the use case above (i.e. TeleSign has no access to call data records (CDRs) of mobile operators).

Due to the number and variety of TeleSign clients, the vast number of mobile phone transactions which pass through TeleSign's systems every day, and the user-initiated nature of these transactions, TeleSign data can be used as an indicator of user activity associated with a given mobile phone number. Detailed activity data direct from mobile operators would clearly give greater population coverage and longer history, but in the absence of availability of such data it is suggested that data sources such as TeleSign can be a good proxy to provide uplift to Verify user approvals.

Objectives

TeleSign developed a prototype real-time service for Identity Providers that returns a response as to whether the mobile phone number has demonstrated activity over a given time period up to 90 days. The service was based on TeleSign's interpretation of the requirements of GPG 45 Element E for Level of Assurance 2 as applied to TeleSign held data. The potential approval uplift as a result of using this service was calculated, as well as a projection of the uplift which could be achieved were mobile operators to provide activity history data directly.

NOTE: Due to restrictions in the open publication of the GDS IPV Operations Manual that contains the scoring metrics for activity event, the following abstraction terminology has been used throughout this document to indicate the segmentation and categorization of data. TeleSign's data was assigned an activity event quality of "Medium" and a longevity of "Short", resulting in a score of "Low". The medium quality assignation is based on there being a robust process in place at the IDP which binds the phone number to the identity over the activity period.

		Longevity of Claimed Identity known by source				
		Very Short	Short	Medium	Long	Very Long
Activity Event Quality	Low	n/a				
	Med	None	Low	Medium	High	Very High
	High	n/a				

Project Phases

The project was divided into 4 phases, with the option to terminate it after phase 2.

Phase 1 – setup - **Complete**

- Create the project scope and detailed plan.
- Ensure that participants have been identified and confirmed as willing to take part.
- Execute initial briefings and confirm the plan is workable.
- Exchange data sharing agreements and NDAs.
- Confirm the size of the sample data list and the mechanism for exchange.

Phase 2 – initial investigation & data quality test - **Complete**
- Participant IDP(s) to submit agreed sample quantity of phone numbers, one batch per IDP.
- TeleSign to check these numbers offline against their database of activity and provide a report on the split between those numbers for which activity is in the database and those where there is no activity.
- For those with activity, the participants will agree how that activity is reflected in the report (i.e. actual activity or an abstracted version thereof).
- Analyse the results and determine whether to proceed to Phase 3.

Phase 3 – build & test API - **Complete**
- Participant IDP(s) and TeleSign build real-time API connectivity to enable IDP(s) to query TeleSign's service.
- API to accept a mobile phone number and return a risk score for that number plus the activity history information for the last 90 days on TeleSign's database.
- Test with fake data at TeleSign's end and confirm technical compatibility and latency.

Phase 4 – test with real data - **Complete**
- When a new registration is being processed and there is insufficient activity information from existing sources, participant IDP(s) trigger real requests into TeleSign and TeleSign responds with real data (risk score + activity history).
- Run for an agreed period to obtain sufficient transactions (over 2000 recommended) to be statistically significant.
- User's actual Verify scores and approval results are not affected by the test. Instead, "shadow" scores are created to demonstrate the impact the TeleSign data could have on the user's approval score were it to be included in the live Verify score calculation.
- Analyse the results and write up the final white paper.

Phase 1 - Setup

After participants signed the appropriate NDA and data sharing agreement, they were free to exchange data for the purpose outlined in the agreement. The approach was as follows:

- Number of mobile phone numbers to be provided from each IDP to TeleSign: 10,000.

- Geographic origin of mobile phone numbers: UK only.

- Format of file: Excel spreadsheet, 1 mobile number per row.

- 1 spreadsheet per IDP. Spreadsheets and results to be kept separate.

- Mobile number format: international country code followed by number, no leading zeroes e.g. 447710761829.

- Excel spreadsheets were encrypted (protect workbook > encrypt with password) and transmitted from each IDP to TeleSign via email, and deleted after use.

NOTE: IDPs included phone numbers of users that didn't make it through the approval process due to a lack of activity history, and those who were successful. This enabled the IDPs to determine if a significant uplift in approval would result from using this new source of activity history data in real time with the approval process. There would be no need to identify to TeleSign whether a number was associated with a successful or failed registration.

Phase 2 -Investigation & Data Quality Test
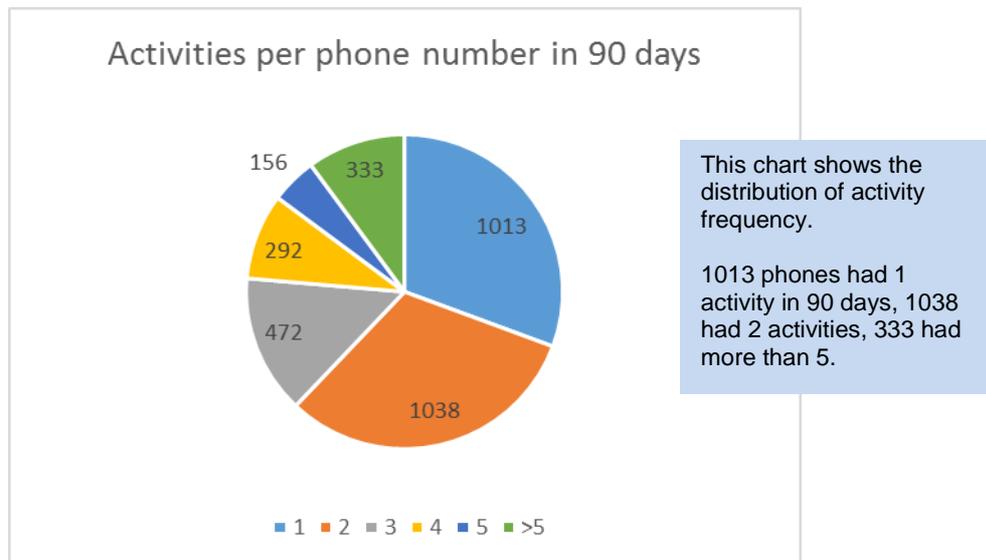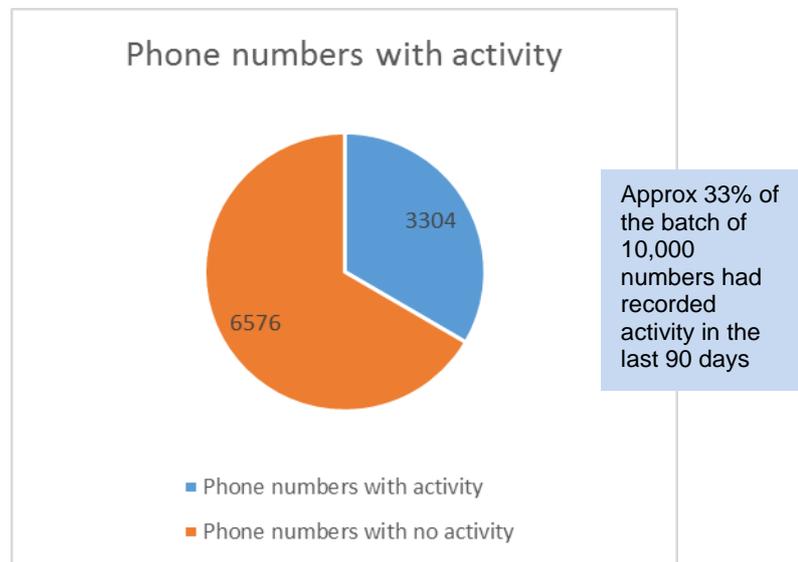
Duration: 3 weeks – COMPLETED.

- The IDPs prepared the data files as per Phase 1 guidelines and transmitted them to TeleSign.

- TeleSign prepared data of recorded activity for UK numbers for the last 90 days.

- For each IDP separately, TeleSign checked each IDP-supplied mobile phone number against TeleSign's recorded activity information and returned a count of the number of times TeleSign had seen that number in its recorded activity data.

- TeleSign prepared a results file per IDP in the form of a spreadsheet containing two columns, the original mobile phone number, and the count of the number of times that phone number was seen at TeleSign.

- TeleSign wrote up an analysis of the results for each IDP.

- TeleSign encrypted each spreadsheet in the same way, and returned them to the relevant IDPs.

- The IDPs each analysed the results and confirmed the implications on their business processes had they been able to get this activity information in real time when processing a new user's registration, or during the periodic re-checking of an existing user's account.

- GDS standards team confirmed the results and the data therein satisfy the requirements of the GPG 45 standard.

- TeleSign wrote up a summary of the data quality test.

Phase 1 and 2 Results Analysis

This is a sample of the data returned by TeleSign to indicate the sort of data returned, showing the total number of activities for a specific number, which is then split across 0-30 days, 31-60 days and 61-90 days.

phonedbphone	delivered	Requests btw 0-30	Requests btw 31-60	Requests btw 61-90
xxxxxxx927	11	3	4	4

The charts below show a high level analysis of one of the IDP's results (both are broadly the same to within 1%), with the split of numbers with and without recorded activity, how that activity is broken down across 90 days of history, and the count of activities per phone within that 90 days.

Phone numbers with activity

3304

6576

Approx 33% of the batch of 10,000 numbers had recorded activity in the last 90 days

- Phone numbers with activity
- Phone numbers with no activity

Activities per phone number in 90 days

156   333

1013

292

472

1038

This chart shows the distribution of activity frequency.

1013 phones had 1 activity in 90 days, 1038 had 2 activities, 333 had more than 5.

- 1 - 2 - 3 - 4 - 5 - >5

The data was provided to Digidentity for analysis. They confirmed that half of the 20,000 numbers that they provided for this trial in the combined sample lists had been declined due to a lack of activity history, and that between 31% and 34% of these could have the possibility of getting a LoA2 account if the TeleSign data were deemed good enough quality from a standards perspective. This is a potential uplift of over 3,000 users within the 20,000 sample.

Discussions with the GDS standards team GDS took place following Digidentity's analysis. GDS confirmed that, for mobile numbers that had been successfully bound to the user by the IDP using a robust process, the TeleSign data quality would be ranked as Medium. Further discussions took place regarding the tenure period (defined as the "weighting of identity events" which related to the "longevity of claimed identity known to source"). TeleSign data for this trial is confirmed to be no longer than 90 days. It was also noted that if tenure information were to be available (the length of time the user has had the mobile number) this could improve the quality rating of the data. For the purposes of this project, tenure was assumed to be up to 90 days as TeleSign has no evidence of tenure beyond that period, enabling a score of "low" to be reached if activity data was present.

The actual uplift of individual users from a fail to a pass is dependent on how close the user was to a pass score. With the maximum additional score that could be obtained is "low", those users within "low" points of a pass could be uplifted from a fail to a pass.

The participants all agreed to proceed to phases 3 and 4.

Phase 3 – API Design, Build, Test

Phase 3 involved designing, building and testing a real time API to obtain the activity history data from TeleSign, as well as a risk score for individual mobile phone numbers. This was used to demonstrate that this data can be obtained quickly enough to be used in flight of a new user registration transaction, and introduce the risk score parameter as a new and potentially useful data point.

TeleSign's API is documented online in the TeleSign Labs repository here:
http://docs.telesignlabs.com/scoreplus_overview.php

For the purposes of this project, the key information returned is in the following blocks of the API:

```
"risk":{
        "level":"medium-low",
        "recommendation":"allow",
        "score":300
    }
    "activity":{
```

Risk score – if TeleSign have records of this mobile number being involved in fraud, this will be high.

Activity counters for user-triggered activity using this mobile phone number on one or more of TeleSign's client systems. Typically 2-factor authentication such as password reset, high value transactions etc.

```
    "requested_activity":{
        "last_30_days":"2",
        "last_31-60_days":"1",
        "last_61-90_days":"4"
    },
 }
```

Additional information in the API such as phone type, carrier name etc is included in the API and may well be useful to the IDPs.

API keys were provided to Digidentity who were able to integrate the API within 2 days and conduct rapid testing before deploying to Live on 21 June 2016.

API Integration & Logic Flow

In order to avoid impacting real user scoring during this Alpha project, a "shadow score" was created by Digidentity for analyzing the results of the TeleSign data. NOTE: The term "score" is used here on the assumption that Digidentity accumulate points into a score to determine if a user is allowed to be provided with a Verify account as per the GDS operations manual. Assumptions have been made as to the exact nature of the Digidentity process flow which remains their intellectual property – Digidentity will need to adapt the logic below accordingly:

The shadow score enabled "before & after" analysis to be done i.e. comparing the actual real result for each user with the result that includes the TeleSign activity history check. This enables an analysis of the potential uplift from fail to pass to be done.

The logic flow was as follows

- Scenario: New applicant registering
- IDP runs through standard flow to the point where the activity history check is inserted.
- At this point the system will check if there is enough Active History
- When there is not enough Active History, create the "shadow score" variable.
- Call TeleSign API, supplying applicant's mobile phone number.
- If result = success
  - Check returned *[risk]* object.
    - If *[riskrecommendation]* = flag or block

- This is a significant contra indicator. It indicates that the number is suspicious. A status of "block" would indicate that the number has been seen in fraudulent circumstances elsewhere. This is a live, real score from Telesign's production database.
- IDP logs the contra indicators without impacting the user's registration process during the Alpha project.

o Check returned *[activity]* object.

  ▪ If any of the counters are > 0

  - Activity has been seen for this number with one or more TeleSign clients.
  - Activity is within the last 90 days, therefore set tenure to be up-to 90 days (for the purposes of this project, this is sufficient).
  - Calculate score as per GDS operations manual
    o data quality = "medium"
    o tenure = "short"
  - Update "shadow score" accordingly.
  - Do not change the actual user score i.e. no impact on the real user's registration journey.

o Once all IDP proofing checks are complete:

  ▪ Write out the actual result of the proofing checks to a file, and write what the result would have been were the TeleSign data to be used. For example, real result = account declined for insufficient evidence. Shadow result = approved for LoA2 account.

  ▪ It is <u>not</u> necessary to log the phone number of the user in the log file.

Live Testing Period and Output

Once integration was complete, Digidentity and Post Office ran the TeleSign API and processed 8784 transactions to obtain a statistically meaningful volume. This exceeded the initial target of 2000 transactions.

Once that live testing period was complete, Digidentity provided an output file which simply confirmed the total number of users for which calls to the API were made, and the number of users who failed to register, but would have been successful if the TeleSign activity information had been used to calculate the user's real score. This is referred to as "uplift". No user identifying information was included in this file.

Phase 4 Results Analysis

At the point where the results file was taken, 8784 transactions had been sent by the participating IDPs through the real time interface. Transactions continued to be sent after this time and over 30,000 have been processed since the API was opened. These 8784 transactions were for users who had inadequate activity information from other sources.

Of those 8784 transactions, 2899 (33%) were found to have activity on the TeleSign database. This is consistent with the initial batch test conducted in Phase 2.

Therefore, 2899 users could have their score increased by the IDP in the specific activity history category. With the TeleSign data being allocated a tenure period 90 days, and being allocated a quality level of "medium", the maximum additional score obtainable under the Operations Manual guidelines is in the category indicated as "low".

Each IDP's specific scoring mechanism is proprietary so the mechanisms of the participant IDPs in this project are not discussed in any detail. Within these constraints, the participant IDPs were able to confirm that approximately 10% of users would have been uplifted from a fail to a pass as a result of this increased score, a total of 291 individuals in this data set. NOTE no actual user scores were changed as part of this project – virtual scores were used instead.

Projected Results for Mobile Operator Data

An additional deliverable for this project was the determination of the potential results if there was activity data available from the UK mobile operators. TeleSign's data has 2 restrictions – firstly it only goes back 90 days due to contractual data retention limits with TeleSign clients. Secondly, it does not cover 100% of the mobile phone using population. Even so, the results of this project indicate it can provide uplift in Verify approval rates.

It can be assumed that UK mobile operator data will cover a much higher proportion of the population, and will have longer data retention periods.

The latter is important as the GDS standards provide more weight to activity that has been carried out for a longer period. Therefore activity data that is older than 90 days enables the IDPs to generate a higher score for an individual application. Therefore, with the combination of greater coverage and longer tenure periods, mobile operator data should provide better results.

In order to provide a calculated guide as to the potential increase in new Verify registration approvals that could result from the use of mobile operator activity data, a calculation has been performed using some simple assumptions.

1) Proportion of adults who personally own and use a mobile phone in the UK: 93% (OFCOM).
2) Data from all 4 main UK mobile operators are included in a potential service.
3) Switching propensity (churn) for mobile 10% per year (OFCOM 2015) though most keep their number, but assume activity starts from scratch with the new operator upon switching. No statistics were found for what proportion of mobile users keep their number when switching, but landline number retention is 74% so we can assume mobile is higher.

4) 95% of UK mobile users will have activity recorded at the 4 main UK mobile operators (GSMA figures). This factors in ignoring the Isle of Man & Channel Islands which have their own mobile operators, and some of the virtual mobile operators who minimise their use of the facilities of the main UK mobile operators. It assumes that mobile operators will be able to obtain activity information for the majority of the larger MVNOs that they provide services for. If operator-hosted MVNOs are not involved (or had not given permission to their hosting MNOs for their customer information to be used), the activity/tenure data coverage falls to 82%.

Therefore assuming average churn of 10%, of 1000 users at a mobile operator 3 years ago, there will be 729 users remaining today, so 729 users will have tenure of 3+ years. There will of course be edge cases whereby a mobile number could be used by more than one person (e.g. husband and wife) and the mobile operators will be able to calculate the effect of these.

Mobile operators will be able to insert their own (commercially very sensitive) churn rates into these calculations, and split prepay, postpay and corporate account types which have different churn rates. These calculations therefore give a guide rather than a definitive prediction.

Using an estimated flat churn rate of 10%, the percentage of users retained by mobile operators over time is as follows (note that the previously described abstraction of the "Longevity of Claimed Identity known by source" from the IPV Operations Manual is used):

MNO retained users	Ops Manual Tenure Category	Score (medium data quality)
<70.4%	Very long	Very high
70.4%	Long	High
88.9%	Medium	Medium
97.5%	Short	Low
98.8%	Very short	None

Using the number of transactions put through the TeleSign API by the IDPs (8784) and multiplying it by the percentage coverage assumption (95%) we can assume that 8345 of the 8784 users would have activity data on the 4 UK mobile operators.

It is then possible to calculate an approximate projection of the number and percentage of user uplift:

Number of users	Score obtainable	Percentage Split
5007	"very high" score	57%
937	"high" score	11%
1475	"medium" score	17%
720	"low" score	8%
8139	total users with increased scores	

93%	of users with increased scores	

When compared to TeleSign data, mobile operator data would yield an additional 5227 users who would obtain an increased score (a 60% increase). Of these 8139 users, 7419 would obtain a higher score than the base 90 day tenure score that TeleSign could return.

For a mobile-operator data source to work as effectively as indicated, the mobile operators would need to implement a similar API to the TeleSign one.. This would involve indicating activity for a given number within the "Longevity of the Claimed Identity" time ranges detailed in the Operations Manual to confirm both activity and tenure. It may be possible for the mobile operators to simply provide a yes/no answer to whether there is activity within those ranges.

Potential Combined Data Approach

For the mobile operators, providing activity information is likely to be more difficult than providing tenure information. This is because activity data will tend to be held as call data records (CDRs) in huge databases. Trawling millions or billions of CDRs in real time to respond to an activity query for an individual mobile number is inefficient and will result in unacceptable latency.

An alternative is to calculate activity counters for all subscribers in the background and put them into a separate database which is easily and rapidly accessible, but this will take time effort for the operators to implement if it doesn't already exist.

Tenure is rather simpler, as a customer record will tend to have a single contract start date or prepay activation date which can be read quickly and efficiently and APIs are likely to already exist to do this). Care will need to be taken to handle the relationship of the mobile number to the contract start date and there will be edge cases that need to be taken into account.

To move rapidly and to increase the efficacy of the TeleSign data tested in this Alpha project, a data combination approach could be used. The TeleSign data only attracts a tenure score of "low" due to the tenure period being "short" i.e. within 90 days. If the mobile operators were able to provide their tenure data, and as seen above that tenure is likely to be significantly longer, and combine it with TeleSign's confirmation of recent activity, it is likely that the 33% of individuals covered by TeleSign's activity could be scored higher due to longer tenure. This will result in an increased percentage of individuals who are uplifted from a fail to a pass. It is recommended that this approach is looked, and GDS approval is sought for the use of combined data of this sort.

Conclusion & Recommendation

TeleSign's data can act as a source for mobile activity information providing there is an acceptance that the 90 day tenure period used in this project is acceptable to the GDS Standards Team to generate the minimum score of "low" and a data quality of "medium", providing that the IDP has linked the users asserted identity details to the mobile telephone number through a robust process. This would also encourage other organisations with near-term activity data to work with the IDPs to plug the gap in available activity information.

The potential of TeleSign's data is to elevate the score in Element E – Activity History of the Claimed Identity, of 33% of new Verify registrants. IDPs will need to calculate the implications on their own transaction scoring metrics to determine the uplift in approvals that could result, using the IPV Operations Manual as a guide. This project has confirmed that an uplift of up to 10% of those users for whom activity data is available, is possible.

Activity data direct from mobile operators can act as a better source of information for the IDPs but is not yet available. The potential is to increase the scores of 93% of Verify registrants who have bound a mobile phone number to their Verify registration information. Due to the longer tenure period that mobile operators have data for, the value of the increased score is likely to be in the higher range of the scores available for this data element. There is excellent potential for mobile operators to help their customers to obtain Verify accounts by making activity data available to IDPs. While data from mobile operators remains unavailable, many customers of mobile operators will be unable to easily and quickly obtain Verify accounts so it is in the interest of the mobile operators to help their customers in this way.

To move rapidly, the potential for combining "relatively easily" obtainable mobile operator tenure data with TeleSign activity data could result in a higher percentage of individuals being uplifted from a fail to a pass, though not as many as if mobile operator activity data were also included. This option should be investigated further in the next stage.

By relying on the IDPs to obtain consent for using the user's mobile number for carrying out such checks, the mobile operators can implement simple API mechanisms which operate in a similar manner to that created by TeleSign without needing to rely on an initial and potentially inaccurate name/address matching transaction. IDPs may want to utilise aggregation services provided in the market that enable them to have a single technical and commercial relationship, rather than separate agreements with each mobile operator.

IDPs will need to analyse the results of this paper to determine how the described service and/or a new mobile operator data service could affect their overall scoring mechanism to determine approval uplift of their individual users.

It is in the interests of the IDPs, the mobile operators, the Cabinet Office, Government online services, and the individuals who are customers of both IDPs and operators, for the UK mobile operators to provide activity and tenure data.

It is recommended that a beta project be initiated as soon as possible to create the required APIs at the mobile operators and IDPs. All parties may wish to engage an aggregator to speed up the technical and commercial process.

The result will be significant opportunities for IDPs to increase the scores they calculate for new users and to therefore increase the percentage of registrants who successfully obtain a Verify account.