messente

SMS Quality

Whitepaper

Uku Tomikas
Lead Sales Researcher

Margus Sütt
Key account manager

Introduction

SMS (Short Message Service) continues to be the predominant message communication platform around the globe. Unlike internet-based messaging services, like Facebook Messenger, WhatsApp, and Viber, SMS is built on open Global Systems for Mobile Communications (GSM) standards, and messages can reach any mobile phone on the planet. Thus, due to the proliferation of mobile devices, application-to-phone (A2P) communication has been on the rise for more than a decade.

In fact, it's expected that A2P traffic will reach over 1.8 trillion messages in 2018, globally.[1] Much of this traffic originates from businesses and organizations communicating with people, whether sending notifications, reminders, marketing messages, or two-way communication. Typically, these organizations utilize an SMS service provider that aggregates mobile networks and their subscribers, allowing

[1] Statista: SMS messages traffic globally

them to reach people across several mobile networks, across multiple geographical regions (also known as the secondary SMS market, as businesses do not go directly to network providers for connections.)

Not all SMS service providers are equivalent, though. There are variances in quality amongst SMS service providers that affect the customer experience and the end-user using the SMS service. While SMS seems simple, there are hundreds of mobile network operators across the globe, allowing SMS service providers to become creative with routing, as well as the technology and logic behind their software.

Defining SMS Quality

The idea of SMS service quality is deeper than the binary notion of delivered or not delivered SMS messages. Beyond the ability to deliver messages, SMS service quality is comprised of four components:

- **Delivery rate** – The ratio of successful deliveries compared the send requests from the software provided by service provider.

- **Latency** – The average amount of time between a send requests and message deliveries.

- **Database quality** – The validity of the mobile phone numbers provided by the client.

- The ability to scale services.

Of course, the components relate to delivery, as the core of any SMS service is message delivery. Each component has its own variables that affect the experience – or quality – of the service.

Delivery Rates: When 1% Makes All the Difference

Delivery rates state the ratio of SMS messages that reach their intended targets, with messages sent as the denominator. It is impossible to achieve 100% delivery rates, as there are several reasons messages may not reach an intended handset. Some causes of undelivered messages are simple, like invalid phone numbers (more on database quality further) and switched off handsets. Others have to do with routing and the effects of multiple "hops" between servers, hubs, and network operators (explained in more detail in section 4.)

Some causes of undelivered messages are simple, like invalid phone numbers and switched off handsets. Others have to do with routing and the effects of multiple "hops" between servers, hubs, and network operators.

Most SMS service providers claim a delivery rate over 90%. While some subjectivism is included when calculating delivery rates, small differences in actual delivery rates have a financially quantifiable result. How SMS service providers manage the variables in the components of their product's quality affects that result.

Most importantly, customers are inclined to select SMS gateway providers solely based on price, which is fair, because the cost of undelivered messages isn't immediately clear. The direct costs of failed SMS messages are the costs of resending messages, if needed, as not all messages have equal importance. Business-critical messages, like notifications and PIN codes, must be resent.

However, there are indirect costs to undelivered SMS messages, whether they are the most critical messages, or marketing messages.

Since most organizations send messages in high volumes, a difference in delivery rate of 1-2% has a significant indirect cost.

The math is straight forward: For an organization that sends 100,000 SMS messages on average, a 1-2% lower delivery rate is 1,000-2,000 undelivered messages. The impact of those undelivered messages depends on the intention of the messages. Further are examples.

SMS Just now
Your PIN code is 2255

you are safely logged in ✓
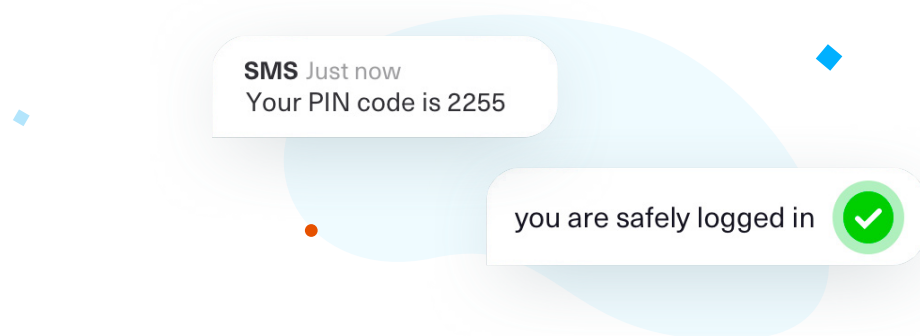
Transactional SMS & PIN Codes
SMS messages used to confirm transactions or PIN codes as a second layer to log into an account (two-factor authentication,) the indirect cost of failed messages is lost revenue. In addition, there's a negative impact on user experience.

Here are some reasons undelivered messages indirectly impact revenue:

- Abandoned account registration process resulting in lost sales & subscriptions

- Cart or sale abandonment when transactions can't be confirmed
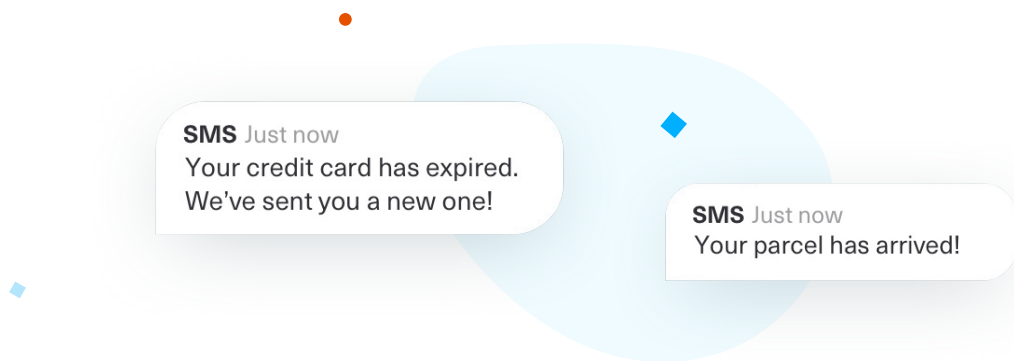
- Lost payment service fees for financial intermediaries

Financial impact to Payment Service Providers
To illustrate, a payment service provider (PSP) that charges (on average) 0.50€ for handling a transaction loses revenue if an undelivered

SMS PIN code results in an uncompleted transaction. In addition, the PSP will likely incur support costs through time spent fixing issues or communicating with the SMS service provider, which vary depending on the quantity of support occurrences (cases or calls.)

Potential financial impact with a 2% lower delivery rate at 100,000 monthly messages:

- 1,000€ of missed transaction fees – 2,000 undelivered PIN codes

- 300-600€ in support time – 15€ an hour for support

SMS Just now
Your credit card has expired.
We've sent you a new one!

SMS Just now
Your parcel has arrived!

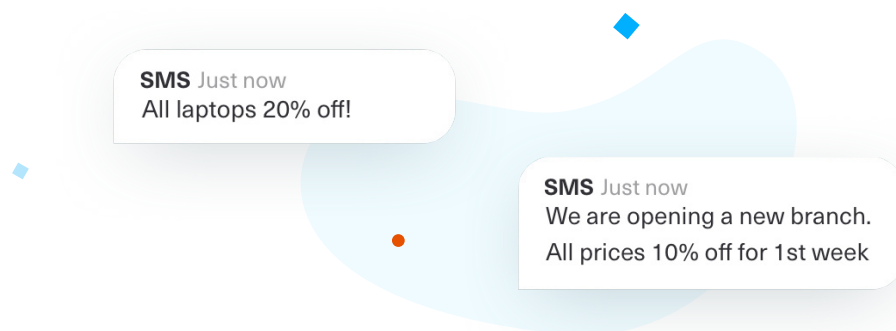**Notifications and reminders**

The indirect costs for missed reminders or notifications vary significantly, as each service using an SMS gateway has varying prices or societal impact; emergency notifications may be life-or-death, while a missed dentist appointment is a loss of 100€ in revenue.

Assuming a dental office uses an SMS gateway to send reminders to 500 patients a month:

- A 2% lower delivery rate results in 10 patients potentially missing their appointment

- At 100€ per appointment, that may lead to a loss of 1,000€ in revenue

Even if the patients reschedule, the revenue is lost at that time, since dentists can't immediately provide that service to someone else –the dental team sits idle.

The same logic applies to any business that loses revenue when the client doesn't show up, from healthcare, to leisure, travel, and hotels.

SMS Just now
All laptops 20% off!

SMS Just now
We are opening a new branch.
All prices 10% off for 1st week

SMS Marketing

Marketers measure results based on conversion rates, so correlating losses is simple.

Let's assume a campaign has a conversion rate of 5%, which is realistic for an SMS marketing campaign, due to high open rates (SMS messages are usually opened within 3 minutes and have a 90% open rate.)[2] Also, the revenue earned on each conversion yields 50€.

The campaign sends SMS messages to 100,000 people:

- Expected revenue at 96% delivery: 240,000€

- Expected revenue at 98% delivery: 245,000€

The difference of 5,000€ pays for the campaign itself.

[2] Tatango: 90% of text messages are read within 3 minutes

Comparing SMS Costs and Direct Losses

The stark difference in pricing between SMS service providers loses its fervor as delivery rates in context create higher losses than even a 30% lower cost per SMS with a 2% delivery rate concession.

The table below illustrates the examples mentioned above with SMS costs included. *The cost per SMS is based on the French market at Messente's list prices.* The cost per SMS at the 96% delivery rate is Messente's price reduced by 30%.

Scenario	Cost at 98% delivery	Cost at 96% delivery	SMS Cost savings	Potential losses
Payment Service Provider	€ 5,300.00	€ 3,710.00	€1,590.00	€ 1,600.00
Notifications & Reminders	€ 26.50	€ 18.55	€ 7.95	€ 1,000.00
SMS Marketing	€ 5,300.00	€ 3,710.00	€ 1,590.00	€ 5,000.00

To conclude that invoice amounts are the only cost to SMS would be financially irresponsible. As each use case results in its own indirect financial impact.

The Effects of Routing and Latency

Global SMS messaging systems come with a level of complexity and complexity creates latency. While a good system sends SMS messages within five to ten seconds, there are several variables that affect latency and may push latency beyond ten seconds.
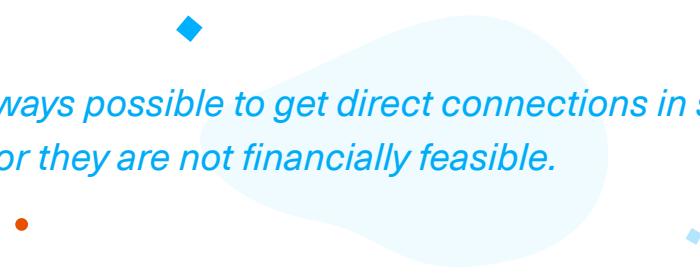
These variables directly affect the quality of A2P SMS messaging, as they are related to delivery, and the experience of end users or intended impact of the message. To illustrate, the effect of a long delay between a send request and delivery is much different with PIN codes

for two-factor authentication than marketing campaigns. PIN codes must be delivered as quickly as possible, as users have limited time before they expire; a marketing SMS delayed by an extra ten seconds will not have an impact, as the desired result is to have recipients read the message and act.

The way SMS service providers manage these variables affects the quality of the A2P SMS service.

Routes and Connections

The amount of connections between the SMS messaging provider and the handsets in question affects the time it takes to deliver messages and throughput.

It's not always possible to get direct connections in some markets, or they are not financially feasible.

Direct-to-network operator connections are ideal, limit latency, and reduce the risk of unexpected latency due to connection issues. Also, direct connections increase control over throughput. However, having only direct connections and creating a truly global network isn't realistic for SMS service providers at the client-side of the value chain; it's not always possible to get direct connections in some markets, or they are not financially feasible.

One-hop connections, where a partner with a direct connection to a specific network operator receives the traffic before it goes to the network operator (then the handsets,) are the next most desired connections. These connections are typically brokered between partners who have a solid relationship and understand the traffic requirements.

Once connections move beyond one-hop connections, the likelihood of increased latency or delivery issues vastly increase, and SMS quality erodes.

Throughput

SMS service providers can commonly accept more traffic than they have in bandwidth from network operators and hubs. This is due to bandwidth being limited by network operators and partners, restricting how much traffic can leave a provider's servers at one time. Direct connections and one-hop connections usually have better throughput. Bandwidth can be negotiated, so strong partner relationships are an important factor for throughput.

Priority

SMS traffic is not equitable. As described earlier, the uses cases for traffic vary and throughput is limited, requiring SMS service providers to define priority. Recall the two-factor authentication PIN code and marketing SMS example: PIN codes must receive a higher traffic priority than marketing messages. Thus, API algorithms must understand when to hold marketing SMS traffic for PIN codes.

Rules and Regulations

Laws vary by market. Regulations can slow traffic, as SMS service providers must be certain that traffic abides by local rules. For example, many countries have non-solicitations laws with varying strictness. Some non-solicitation laws require traffic to be audited against non-solicitation lists before SMS messages are sent to handsets. However, if the message is not a marketing message, it does not need to be checked against the non-solicitation database, building on the need for the service provider to be able to understand the SMS traffic.

Physical Limitations

Physical limitations are the least controllable variables and are worth mentioning. It's important to understand that that location of the servers making the API requests in relation to the SMS service provider's servers, and network hubs, and mobile operators, will make a difference. While the latency changes are fractions of seconds, the data must travel through a network before it reaches the handset.

Database Quality

Delivery rates are calculated based on successful deliveries compared to attempted SMS deliveries to valid mobile phone numbers. However, the quality of a client's mobile phone number database will impact the delivery rate for that specific client. In addition, clients are charged for each message *sent* regardless of whether it is delivered.

What Effects Database Quality

Mobile phone numbers are provided by the end customer. Thus, the validity of the phone numbers is reliant on the clients' customers and the industry in which the client operates in is an indicator of potential invalid, fake, or virtual mobile phone numbers. To illustrate, payday loan companies tend to have significantly more invalid or fake phone numbers. On the other hand, banking customers are not likely to provide invalid numbers, as they use their mobile phones to receive PIN codes for two-factor authentication and transaction confirmations.

Messente's experience shows that invalid and roaming mobile numbers amount to 8-11% of client databases.

In addition, people change their phone numbers, use temporary numbers (temporary SIM cards,) or cancel their service. Ultimately, clients should not assume that mobile numbers that they have collected are completely valid. Messente's experience shows that invalid and roaming mobile numbers amount to 8-11% of client databases.

Database Quality Solution

It is best practice to check the validity of the phone numbers in a database before engaging in communication. Primarily, removing invalid mobile numbers reduces cost, as clients avoid attempting to send messages to mobile numbers that do not work. With a clean database client specific delivery rates improve.

Messente provides an additional tool that allows clients to scrub their databases before moving forward with sending messages. The number lookup API provides pertinent information to clients, giving them more insight on the mobile phone numbers they've collected.

The number lookup API pings each phone number:

01   To check if it is valid

02   If it is roaming

03   Get information about which network operator it belongs to

04   And if the number has been ported from another operator

Users of Messente's services can use this API at their discretion. The cost to ping a number is much less than sending a message and end users do not know if their number was pinged. The frequency of running the API depends on the SMS use case and how often the phone number database changes.

Note: Some network operators and some countries do not provide delivery reports. Delivery rates and number lookup are contingent on whether the operator or country allow for delivery reports.

The Final Variable: Scaling

The details in this report consider service quality. Yet achieving SMS service quality at scale is more relevant than the details of the quality itself. Businesses looking for a global SMS service provider are typical multinational companies or plan to scale to multiple regions in the world.

messente

Scaling the service includes all aspects of the service. Barring limitations from mobile network operators, country specific restriction and legislation, quality SMS gateway providers have the capability to provide the same service across the world. This includes support, routing, and database lookups.

Hence, Messente works closely with partners (network operators and hubs) to ensure that all services offered follow the client. If Messente doesn't cover a region with all its services, the team will work to make it happen.

References
SMS quality report

Statista GmbH. "A2P & P2A SMS traffic worldwide by region 2011-2018 | Statistic." Last modified 2018. Accessed 18 Jan 2018. https://www.statista.com/statistics/485141/a2p-and-p2a-sms-traffic-worldwide-by-region/.

Tatango, Inc. "90% of Text Messages Are Read Within 3 Minutes | Tatango." Last modified 2012. Accessed 21 Feb 2018. https://www.tatango.com/blog/90-of-text-messages-are-read-within-3-minutes/.