**Etude de la relation entre churn et comportement RGPD dans les télécommunications: une analyse par les tables de survie**

Eric Di Benedetto

Doctorant, CEMOI, IAE de la Réunion

eric.di-benedetto@univ-reunion.fr

Alain Cucchi

Professeur, CEMOI, IAE de la Réunion

alain.cucchi@univ-reunion.fr

**Résumé**

Ce papier de recherche s’inscrit dans l’étude du comportement des consommateurs dans le secteur des télécommunications. La majorité des recherches en datamining, en particulier chez les opérateurs, est dédiées à la fidélité du client. Le churn est un indicateur clé particulièrement suivi dans l’étude de la performance de l’entreprise et de la fidélité client. Par ailleurs, la littérature fait état d’une confiance supérieure du client et de résultats améliorés lorsque les données personnelles des consommateurs semblent utilisées à bon escient. Nous proposons d’explorer la relation entre le churn des clients et leur comportement d’exercice des droits sur leurs données personnelles. Les premières analyses datamining tendent à prouver que les clients qui utilisent leurs droits RGPD sont plus fidèles et leur ancienneté, plus élevée.

Mots-clés : analyse de survie, churn, rétention, RGPD, ancienneté client, télécommunications, vie privée

**A study of churn and GDPR behaviour in the telecommunication industry: a survival analysis**

**Abstract**

This research paper is part of a larger study of consumer behaviour in the telecommunications industry. Most of the research in datamining, especially among operators, is dedicated to customer loyalty. Churn is a key indicator particularly monitored in the study of company performance and customer loyalty. Moreover, the literature reports superior customer trust and improved results when consumers' personal data appear to be used appropriately. We propose to explore the relationship between customer churn and their behaviour in exercising rights over their personal data. Initial datamining analyses tend to show that customers who use their GDPR rights are more loyal and have a higher level of seniority.

Keywords : churn, GDPR, privacy, telecommunications, customer seniority, survival analysis

**Introduction**

Le secteur des services, et des télécommunications en particulier, connaît des mutations technologiques rapides, notamment avec l’évolution vers la norme 5G, le déploiement de la fibre dans les territoires, pour s’adapter à l’explosion des usages. Mature, le marché est dominé par une quarantaine d’opérateurs d’envergure mondiale (ITU, 2019) qui revendiquent à eux seuls près de 9 clients sur 10 dans le monde (ITU, 2019). D’accords commerciaux en restructurations, les opérateurs sont en permanence à la recherche d’avantages concurrentiels qui leur permettraient d’attirer de nouveaux clients, à l’image de l’intégration de la domotique, de services de presse en ligne, ou de vidéo à la demande. C’est pourquoi les fournisseurs de services de télécommunications optimisent en permanence leurs offres, et proposent, pratiquement chaque mois, de nouvelles promotions adaptées aux cibles qu’elles souhaitent atteindre. Le client souhaitant toujours plus de service pour un prix toujours moindre (Lu, 2002), il est primordial d’analyser sa demande, pour y répondre avec exactitude et si possible, l’anticiper, par exemple en analysant les réponses aux sollicitations dont il fait l’objet. Utilisées à bon escient, ces demandes relatives aux données personnelles permettraient d’augmenter la confiance du consommateur, tout en diminuant la vulnérabilité perçue (Deslée, Collin-Lachaud et Diallo, 2018). Selon Ngai et al. (2009), ces techniques d’analyse de données « permettent de prédire les comportements des clients », en accélérant et fiabilisant les processus de décision. Selon la méta analyse de ces auteurs, ces études en datamining sont essentielles à la fidélisation des clients, à tel point qu’elles consacrent plus de 60% des recherches. Ainsi, le Management de la Relation Client (CRM) est à la recherche d’efficience pour optimiser ses coûts et répondre à la demande (Lu, 2002 ; Larivière, Van den Poel, 2004).

Cependant, les opérateurs ont dû s’adapter aux garde-fous mis en place par les Etats pour protéger leurs marchés face à l’ouverture du secteur, mais aussi aux règles de protection du consommateur, notamment face aux enjeux que pose la protection de la vie privée et des données personnelles. Le Privacy Shield (2016), et le Règlement Général sur la Protection des Données (RGPD – 2018) constituent désormais le socle de la protection des données à caractère personnel des citoyens Européens, en réponse notamment aux abus et fuites de données qui ont émaillé le secteur des services au cours de la dernière décennie. Ce papier de recherche s’inscrit dans le cadre d’une recherche plus large sur la relation entre rétention des clients et la protection des données personnelles -RGPD- dans le secteur des télécommunications. Par l’analyse de l’appropriation de leurs données personnelles par les clients (dites « demandes RGPD »), nous recherchons si celle-ci a une influence sur le churn, et par conséquent, la performance de l’entreprise.

**Revue de littérature**

**Le churn, une préoccupation majeure pour les opérateurs de télécommunication**

Dans ce contexte particulièrement contraint et concurrentiel, la rétention des clients a fait l’objet de nombreuses études. Si les chiffres du churn varient selon les pays, les secteurs et le type de client (Hung et al., 2006 ; Lu, 2002), il est communément admis depuis plusieurs décennies que l’acquisition d’un nouveau client coûte généralement 5 à 6 fois plus cher que sa rétention (Rosenberg, Czepiel, 1984). C’est pourquoi l’attrition des clients est devenue un enjeu majeur, si ce n’est de survie dans le secteur des télécommunications (Hung et al., 2006). Ainsi, le « churn management » vise à fidéliser les clients à haute valeur (Berson et al., 2000), et permettrait « d’augmenter les profits de 85%, simplement en réduisant le taux d’attrition de 5% » (Larivière, Van den Poel, 2004). Soumis à des évolutions rapides de technologies et de normes, les firmes ont su rapidement se tourner vers le datamining pour enrayer ce phénomène de churn (Hung et al., 2006). Si certains facteurs bien connus tels que l’âge ou l’ancienneté semblent favoriser le churn (Kim et al., 2004 ; Hung et al., 2006), l’évolution normative a permis, en outre, d’en diminuer le coût pour le consommateur, en France avec la loi Chatel de 2008. L’effet « barrière » des coûts de churn (Kim et al., 2004), tend à diminuer, grâce à la concurrence, et au taux important d’équipement des ménages dans le monde (ITU, 2019).

**Le RGPD et les demandes relatives à la vie privée**

Si la privacy fait l’objet d’une littérature pléthorique et de méta analyses, elles-mêmes nombreuses, le RGPD est récent. Il semble avoir profondément modifié le marketing digital avec l’apparition, obligatoire depuis le 25 mai 2018, des bandeaux de consentements aux cookies, mais aussi la présence de mentions légales relatives à l’acceptation des choix de gestion de la vie privée. Depuis deux ans, le client a donc la possibilité, dorénavant sous peine de sanctions de la CNIL, d’obtenir l’accès -notamment- à ses données personnelles détenues par l’entreprise, de s’opposer à toute opération de marketing direct ou indirect, mais aussi de faire l’objet d’études de type datamining. La loi informatique et libertés de 1978 a donc été considérablement renforcée, et le contrôle du consommateur sur sa vie privée, étendu. La recherche a donc exploré ce nouveau champ d’études, notamment sous les prismes juridiques et financiers. Le marketing s’est quant à lui logiquement orienté vers le numérique, explorant les possibilités d’anonymisation des donnes, le big data, ou de mise en confiance du consommateur (Martin et al., 2017), de transparence (van de Waerdt, 2020), sur ces sujets épineux, et dorénavant règlementaires, de vie privée.

Depuis l’instauration de cette loi, il est impératif de tenir compte des consentements et choix du consommateur relatifs à la gestion de sa vie privée. Aussi, il devient nécessaire de mettre en œuvre des techniques de purge des données avec des dates d’expiration relatives aux finalités de traitements (article 5 du RGPD), mais aussi de respecter le droit du consommateur à limiter ou s’opposer aux traitements automatisés (articles 18 et 21 du RGPD). Des mesures techniques et organisationnelles ont été prises pour s’adapter à ces contraintes. Citons par exemple l’obligation de nommer un Data Protection Officer (DPO, article 37 du RGPD), dès lors que l’entreprise traite des données personnelles à grande échelle, et la nécessaire mise en place du recueil des consentements des consommateurs. Ceux-ci sont recueillis dans leur espace personnel (le « selfcare »), mais aussi par email, par le service client au téléphone, par courrier, ou par SMS dans le cas du « Stop SMS ».

**Description des données et méthodologie**

Nous avons souhaité analyser le comportement d’attrition des clients d’un opérateur Français de télécommunications en outremer. Notre analyse longitudinale s’étend sur 24 mois. Elle débute au mois de mai 2018, à l’avènement du RGPD. 851 042 clients actifs sur une durée minimale d’un mois au cours de cette période ont été étudiés. La base de données marketing de l’étude comprend les données issues de l’outil de CRM de l’opérateur, mais aussi celles liées aux usages des consommateurs, ainsi que leurs choix relatifs à l’utilisation de leurs données personnelles. Les consentements des consommateurs se répartissent en deux catégories principales:

* Stop au démarchage publicitaire (« Opt-out »). Le client a choisi de ne plus recevoir de sollicitations publicitaires par SMS, de newsletters, de courriers, ou de démarchages téléphoniques. Ces demandes non exclusives, peuvent être cumulées.
* Stop analyse marketing (« Opt-in »). Le client a souhaité ne pas faire l’objet d’études marketing. Auquel cas, il est exclu de notre recherche pour respecter la législation en vigueur.

Les données ont été anonymisées en respectant autant que possible les principes du Privacy Preserving Datamining (Mendes, Vilela, 2017), et leur conformité au RGPD contrôlée par un DPO. Enfin, nous avons procédé au « minage » des données, dans le but d’étudier la relation entre RGPD et churn.

**Résultats**

Le « minage » des données, mené itérativement, a débuté dans une optique exploratoire. Il était primordial de réaliser une segmentation des clients, pour monter en compétence sur la connaissance client, afin de poursuivre l’étude. Les consommateurs ont été clusterisés par âge, genre, ancienneté, type et volumes d’usages consommés et facturés, ainsi que par produits souscrits. Chacune des 30 variables dont fait l’objet cette étude -sur les 300 dont nous disposons- a été sélectionnée au regard de sa présence dans la littérature sur le churn. Les régressions pratiquées, en confirmant des résultats « classiques » de la littérature (Kim et al., 2004 ; Hung et al., 2006), ont permis de valider la sélection des variables pour cette étude initiale. Si nous constatons, sans surprise, que les clients qui procèdent à une demande RGPD sont sur représentés parmi ceux disposant d’une offre simple, en ligne, le corolaire est en revanche un revenu moyen par utilisateur inférieur de 20%.

7 988 clients ont effectué une demande RGPD relative au démarchage publicitaire. Si ce chiffre semble faible (0,94% des clients sur la période), il traduit des comportements intéressants. 34% d’entre eux se sont ainsi opposés à toute forme de démarchage publicitaire, 43% d’entre eux s’opposant à un seul média de communication. Parmi ceux-ci, 2 714 clients se sont opposés aux publicités par SMS. Les publicités par « texto » semblent concentrer les principaux griefs des clients, puisqu’elles représentent 6 754 demandes RGPD, soit plus de 54% des demandes.

Cette première phase d’étude passée, nous avons été frappés par l’ancienneté moyenne des clients RGPD, qui nous semblait importante. Nous avons alors constaté une relation significative (p=0,000), bien que relativement faible, entre le churn et les demandes RGPD.

En observant les résidus, nous constatons que les clients ayant effectué une demande RGPD relative au démarchage publicitaire sont significativement moins présents parmi les churneurs. Les clients qui usent de leurs droits associés au RGPD ont donc moins tendance à quitter l’opérateur.

Pour nous assurer de l’exactitude de cette affirmation, nous avons procédé à une analyse de survie (Ngai et al., 2009 ; Lu, 2002) par l’estimateur de Kaplan-Meier.

Le choix des abscisses et ordonnées de notre étude est dicté par plusieurs considérations, d’ordre contextuel, et de lisibilité.

 \_La première considération est liée à la forte saisonnalité du churn et des conquêtes clients. Sur la période de 24 mois étudiée, nous constatons, chaque année, une forte hausse du churn en fin d’année scolaire (mai-juin), qui s’explique par le nombre important de personnes quittant le territoire: étudiants et mutations professionnelles. A l’inverse, les mois d’août et de décembre sont marqués par des temps forts marketing (« back to school » et fêtes de fin d’années) au cours desquels les opérateurs communiquent fortement. Le churn diminue fortement, les conquêtes de clients augmentent, mais du fait de la publicité, les demandes RGPD sont en hausse. Par ailleurs, outre l’avènement du RGPD en mai 2018, deux phénomènes ont eu une incidence sur le churn et les demandes d’exercice des droits personnels. Le premier est le grand évènement sportif de juin-juillet 2018, sur lequel les opérateurs ont particulièrement communiqué, augmentant les demandes RGPD, et la conquête de nouveau clients. Le deuxième phénomène visible constaté dans notre étude est celui lié à la crise sanitaire mondiale du COVID-19. En mars-avril 2020, le churn et les conquêtes ont été pratiquement nuls. Cependant, nous avons constaté une forte hausse des demandes liées aux données personnelles lors des 3 premières semaines du confinement. Disposant de plus de temps en ligne (augmentation de 25% du trafic internet sur la période), ils auraient davantage contrôlé leurs données et l’opérateur, et testé sa transparence (Martin et al, 2017). Partant de cette première considération, un graphique temporel sur 24 mois constitué de pics importants, difficilement lisibles, nous paraissait devoir faire l’objet d’un papier dédié, en segmentant les clients.

\_Enfin, le choix de l’ancienneté client en abscisse nous permet de préserver de conserver plus aisément la confidentialité du churn, sur lequel les opérateurs de services Européens communiquent rarement.



Figure 1. Analyse de survie des clients par ancienneté selon leur nombre de demandes RGPD.

L’étude des tables de survie et du graphique en courbes associé montre clairement que les clients ayant pratiqué le refus de sollicitations commerciales churnent moins que les clients n’ayant pas opté pour l’exercice de leurs droits RGPD. Au cours des 24 mois étudiés dans ce papier, les clients ayant effectué 4 demandes RGPD (Stop SMS, Stop newsletter, Stop courrier, Stop télévente), ont une ancienneté significativement supérieure aux autres clients, et ce faisant, sont plus fidèles. Selon nous, ces clients, attachés aux valeurs de respect de la vie privée s’expriment auprès de l’opérateur en exerçant leurs droits. La réception de sollicitations commerciales, non désirées par la pratique de l’optout, rendent le contrôle à l’utilisateur (Martin, et al. 2017).

**Discussion, limites, et voies de recherches**

Notre étude s’inscrit dans le cadre d’une recherche plus globale sur le « comportement RGPD » du consommateur, et du churn en particulier. Il s’agit à notre sens de la première du genre, sur un sujet très actuel, et qui reste encore à explorer. Cette recherche nous semble prometteuse. Menée depuis un an, et dans une optique initiale de datamining marketing, le comportement des consommateurs étudiés s’affirme. Il conforte chaque jour les premiers résultats. Par ailleurs, le SMS semble concentrer la majorité des refus publicitaires des clients. Si le « STOP SMS » peut s’exprimer par SMS, ou via l’espace personnel du client, il fait l’objet de nombreux griefs lorsqu’il est exprimé par téléphone avec le service client, et semble vécu comme étant particulièrement intrusif. Si « la mauvaise utilisation des données personnelles n’implique toutefois pas toujours une fidélité comportementale » (Deslée, Collin-Lachaud et Diallo, 2018), il semble nécessaire d’améliorer la compréhension de ce qu’est une « mauvaise utilisation », notamment au regard des résultats des « STOP SMS ». L’injonction aux entreprises de proposer des pages de personnalisation des données de l’utilisateur aurait un impact bénéfique sur le Privacy-Personnalisation Paradox (Awad et Krishnan, 2006). En parallèle, d’autres données sont explorées, qui nous permettraient de mieux expliquer ce comportement client, tout en explorant différentes techniques de datamining (Ngai et al., 2009 ; Hung et al., 2006). L’intérêt managérial est triple. D’une part, il permet d’évaluer la mise en conformité RGPD de l’opérateur, et d’améliorer l’exploitation des données personnelles. D’autre part, il permet de déterminer à travers le temps, l’efficacité et l’impact des campagnes marketing menées à travers le temps. Enfin, il permet d’améliorer la connaissance du client, et ce faisant, de lui répondre plus efficacement. Cette recherche souffre en revanche de limites. La première tient en l’absence de l’exploitation de techniques récentes de datamining de type big data ou deep learning, la raison principale tenant en la relative faiblesse de volumétrie client ne permettant pas d’exploiter au mieux ces techniques. Cette voie devra être explorée lorsqu’il s’agira d’intégrer les 270 variables dont nous disposons, mais non encore intégrées. La deuxième limite tient en la difficulté règlementaire de pratiquer des entretiens qualitatifs avec les clients. Si le RGPD permet désormais au client de refuser d’être contacté par son opérateur à des fins publicitaires, ou d’exploitation de ses données, il paraît difficile de lui demander de participer à un entretien, et la question est éthiquement discutable. La théorie du Privacy Communication Management Theory (Petronio, 2002) nous semble être une piste à explorer. Selon cette théorie, le consommateur définirait la limite entre l’information privée et celle, publique, qui peut être exploitée et divulguée par l’opérateur. Tant que la « limite de turbulence » n’est pas franchie par le fournisseur de services; notamment en termes d’envois publicitaires; le client agirait en confiance, sa propension à churner diminuant en conséquence. Par ailleurs, il semblerait que la prise de contrôle volontaire (optout) du client face à l’exploitation de ses données personnelles marque le souhait de poursuivre la relation dans la durée avec l’opérateur. Ce faisant, le RGPD instaurerait peut-être plus qu’une relation de contrôle entre le client et la marque. Il pourrait s’agir d’une relation de confiance mûrement choisie, qui modifierait le comportement du consommateur (Martin et al., 2017).

**Références bibliographiques**

Awad, N. F., & Krishnan, M. S. (2006). The personalization privacy paradox: an empirical evaluation of information transparency and the willingness to be profiled online for personalization. *MIS Quarterly,* 13-28.

Berson, Alex, Smith, Stephen and Thearling, K. (2000) Building Data Mining Applications for CRM, *McGraw-Hill*, New York, NY, 2000.

Deslée A., Collin-Lachaud I. et Diallo M.F. (2018), Les conséquences de la divulgation des données personnelles sur le vécu de l’expérience online du consommateur, papier présenté au Congrès de l’Association Française de Marketing, Strasbourg (16-18 mai), France.

Hung, S. Y., Yen, D. C., & Wang, H. Y. (2006). Applying data mining to telecom churn management. *Expert Systems with Applications*, 31(3), 515-524.

International Telecommunications Union (2019, décembre 20). Statistics. *ITU*. <https://www.itu.int/en/ITU-D/Statistics/Pages/stat/default.aspx> (accédé le 2 juin 2020).

Kim, H.-S., Yoon, C.-H. (2004). Determinants of subscriber churn and customer loyalty in the Korean mobile telephony market. *Telecommunications Policy*, *28*(9‑10), 751‑765.

Larivière, B., & Van den Poel, D. (2004). Investigating the role of product features in preventing customer churn, by using survival analysis and choice modeling: The case of financial services. *Expert Systems with Applications*, 27(2), 277-285.

Lu, J. (2002). Predicting customer churn in the telecommunications industry––An application of survival analysis modeling using SAS. *SAS User Group International* (SUGI27) Online Proceedings, 114-27.

Martin, K., Borah, A., Palmatier, R. (2017) Data privacy: Effects on customer and firm performance. *Journal of Marketing, 2017,* vol. 81, no 1, p. 36-58.

Mendes, R., Vilela, J. P. (2017). Privacy-Preserving Data Mining: Methods, Metrics, and Applications. IEEE Access, 5, 10562–10582.

Ngai, E. W. T., Xiu, L., & Chau, D. C. K. (2009). Application of data mining techniques in customer relationship management : A literature review and classification. *Expert Systems with Applications,* 36(2), 2592‑2602.

Petronio, S. (2002). Boundaries of privacy. *New York: State University of New York Press.*

Rosenberg, L. J., & Czepiel, J. A. (1984). A marketing approach for customer retention. *Journal of consumer marketing*.

Van de Waerdt, P. J. (2020). Information asymmetries: recognizing the limits of the GDPR on the data-driven market. *Computer Law & Security Review,*38, 105436.