Subject: **Fwd: Response for consultation paper on Unsolicited Commercial Communication**

Date: 11/10/17 12:25 PM

From: "Asit Kadayan, Advisor" <advqos@trai.gov.in>

To: rajender@trai.gov.in, pkg20672@gmail.com, trishna.dot@gmail.com

Consultation Paper on Unsolicited Commercial Commu... (141kB)

-------- Original Message --------
From: **Subir Mansukhani** <subir.m@gmail.com>
Date: Nov 10, 2017 10:59:33 AM
Subject: Response for consultation paper on Unsolicited Commercial Communication
To: advqos@trai.gov.in

Dear Shri. Asit Kadayan,

Please find attached a word document that has some responses to the questions that was put out for consultation in the consultation paper on Unsolicited Commercial Communication.I hope the responses are useful; we are always glad to help out on the ML/AI aspects.

---

Regards,
Subir

Q. 1. To what extent, time required for registration and enforcement can be reduced? For achieving reduced time lines, what changes in processes or in different entities e.g. PCPR, NCPR, CPDB may be required? Will providing scrubbing as a service for RTM reduces time? Please give your suggestions with reasons.

None of the SLA should exceed a day. Within a day batches can be run to process all requests. Scrubbing as a service is definitely helpful as it can then be built into a process flow that is completely automated with latest information being available at all times

Q. 2. How to ensure availability of Mobile Apps for registering preferences and complaints and for de-registration for all types of devices, operating systems and platforms? Whether white label TRAI Mobile App may be bundled along with other Apps or pre-installed with mobile devices for increasing penetration of app? For popularizing this app, what other initiatives can be taken? Please give your suggestions with reasons.

It's better to let the TSPs build the app interfaces to register the preferences and complaints. APIs should be able for the TSPs to submit the data to the server side. A sample white label app/apk can be provided that the TSP can use in case they don't want to build such capability into their apps.

Q. 3. In case of Mobile Number Portability (MNP), what process may be defined for retaining the status of customer for preference registration? Please give your suggestions with reasons.

The preferences can simply be transferred. The preferences can anyway be updated, on moving to a different TSP an SMS can be sent saying that previous preferences have been transferred and can be updated later

Q. 4. How bulk registration may be allowed and what may be the process and documents to register in bulk on behalf of an organization or family? Please give your suggestions with reasons.

Preferences are typically individualistic in nature.There are concerns if bulk registration takes place without user consent. An entity might be able to bulk register many users and then target them with messages if the process of bulk registration is not fool proof."

Q. 5. Is there a need to have more granularity in the choices to actually capture customers interest and additional dimensions of preferences like type of day, me- 21 dia type(s)? What will be impact of additional choices of preferences on various entities like CPRF, PCPR, NCPR, CPDB etc.? Please give your suggestions with reasons.

There is a need both from a customer convenience standpoint and an enforcement standpoint.However this would need that some rules have to be evaluated at the TSP or RTM end to ensure that these are not being violated and if done in large numbers can raise the amount of computation required to check the rules.

Q. 6. Should the scope of UCC regulation be enhanced to include unwanted calls like silent, obnoxious, threatening calls etc. and unauthorized communications.? What role government

or constitutional organizations may play in curbing such activities? Please give your suggestions with reasons.

This should be enhanced with severe punishments for such violations.Consumers should be able to report such violations to government organizations and the details of the follow up status should be made available to the reportee.

Q. 7. What steps may be taken to address the issues arising from robo-calls and silent calls? What are the technical solutions available to deal with the issue? How international co-operation and collaboration may be helpful to address the issue? Please give your suggestions with reasons.

Advanced big data call graph analytics would be useful in this area. However the solution is not trivial from an engineering and financial point of view. The investment would need to be put in by the TSP just for the purpose of preventing such issues which might not be on top of their agenda as it doesn't drive revenues.Some support could be extended from the DoT from the fines that collected from the violators.

Q. 9. Should registration of other entities such as content providers, TMSEs, Principal Entities, or any other intermediaries be initiated to bring more effectiveness? Whether standard agreements can be specified for different entities to be entered into for playing any role in the chain? Please give your suggestions

Registration of all entities along the chain should be mandatory. The entities have already be defined and they should all be KYCed/eKYCed. Every parent in the chain should be held responsible for the immediate child in the tree.Standard agreements can be and a list of acceptable documents can be prepared. The ASA/AUA/subAUA model that UIDAI follows can be looked at for some ideas as well.

Q. 12. Whether scrubbing as a service model may be helpful for protection of NCPR data? Whether OTP based authentication for queries made by individuals on NCPR portal may be helpful to protect NCPR data? What other mechanisms may be adopted to protect the data? Please give your suggestions with reasons.

Scrubbing as a service is a good idea. Based on key exchanges etc. the service APIs can be controlled only to authenticated users and its easy to audit as well since there is a trail of every call made to the service.OTP based authentication is good but not fool proof. The keys can be embedded into a calling application and encrypted data can be transferred to the calling application which decrypts it. Additionally periodic audits should be conducted at places where lots of requests seem to be originating and by using analytics to detect similar sources that are suspicious.

Q. 16. What steps need to be initiated to restore the sanctity of transactional SMS? What framework need to be prescribed for those transactional SMS which are not critical in nature? Please give your suggestions with reasons?

All messages/message templates should be approved before making them live. Some inspiration can be taken from the way Google operates to list sites on its advertising platform. First a machine looks at the content and determines all the criteria that is laid out before hand, this involves Machine Learning algorithms as content is free form and cannot be templatized completely. If it is not met then the process takes a little more time for approval by getting human involvement. All approved messages and their headers should sit in a centralized database and can be checked in real time. The centralized approved message database can be jointly maintained by all the TSPs.

Q. 19. Whether access providers may be asked to entertain complaints from customers who have not registered with NCPR in certain cases like UCC from UTM, promotional commercial communication beyond specified timings, fraudulent type of messages or calls etc.? What mechanism may be adopted to avoid promotional commercial communication during roaming or call forwarding cases? Please give your suggestions with reasons.

APs should record complaints from customers who are not registered with NCPR and these customers should be made aware of NCPR and be asked if they would like to be put in the NCPR database along with their preferences.For preventing cases of UCC listed above, in addition to a first filter based on rules a Machine Learning solution should be put in place at the AP that is data driven.Data should be tagged on an ongoing basis so the ML algorithms keep getting better over time. Constantly updating rules will only result in a system that is lagging behind real world situations, instead the solution should use ML and learn as and when more data is collected.

Q. 20. How the mobile App may be developed or enhanced for submitting complaints in an intelligent and intuitive manner? How to ensure that the required permissions from device operating systems or platforms are available to the mobile app to properly function? Please give your suggestions with reasons.

The mobile app should present the user with messages it thinks as spam and ask the user to confirm if the messages it think as spam are indeed spam. Additionally it should let the user mark and submit messages as spam that the app has not identified.The app should ask for permission to read the users SMS and address book,however there are issues to do this in the case of iOS devices.The system to detect spam in SMS should be a combination of both a rule based system and a machine learning algorithm. There should also be a server side component of this algorithm that runs this message and other available aggregate "features" computed for the originator against a model that has been built on messages that were actioned for penalty by the TSP to verify that the message is indeed spam.

Q. 21. Should the present structure of financial disincentive applicable for access providers be reviewed in case where timely and appropriate action was taken by OAP? What additional measures may be prescribed for Access Providers to mitigate UCC problem? Please give your suggestions with reasons.

The present financial disincentive should be applicable. The OAP should be responsible for content that is being sent using their network. Every message should be scored in real time for content that doesn't comply with the rules. With current Machine Learning algorithms.

very accurate solutions can be built and messages can be checked in a matter of milliseconds if they are violating norms.

Q. 23. What enhancements can be done in signature solutions ? What mechanism has to be established to share information among access providers for continuous evolution of signatures, rules, criteria?Please give your suggestions with reason.

Signature solutions should employ newer algorithms instead of just relying on rule based systems like throughput rate of messages, time etc. They should take into account call graphs, message content, meta data associated with calls and SMS's at a macro network level as well at a micro level which is content,originator, duration of call etc. APs can form a consortium and maintain a central database of content and originators that violate norms. Once access to data is democratized then many different algorithms can be built on top of the data.

Q. 24. How Artificial Intelligence (AI) can be used to improve performance of signature solution and detect newer UCC messages created by tweaking the content? Please give your suggestions with reasons.

Big Data and AI should be used in conjunction with each other to build a more robust signature detection solution. ML/AI algorithms can be used to detect spam messages, spam calls etc. at a micro level and Graph analytics can be used to detect sources of origination of spam messages , spam calls etc. Signatures can be built based on data as well as meta data of the information that is being collected.ML/AI solutions are robust to changes in content as they are not rule based and learn patterns in a more flexible way than traditional systems. Cloud would be a crucial component here as well because of the sporadic high compute requirements of AI/ML. These solutions however require sophistication from an engineering and mathematical point of view and can turn out to be expensive to implement the first time around. The crux of ML and AI is data, so a good solution can be built as long there is access to a lot of data to train the algorithms.Central repositories without PII data can be setup so data amongst TSPs can be shared and that their solutions can use to implement and put in place better algorithms.

Q. 26. Should the data from mobile app or from any other source for registering complaints be analyzed at central locations to develop intelligence through crowd sourcing? How actions against such defaulters be expedited?Please give your suggestions with reasons.

Registered complaint data should be stored and analysed at a central location. New patterns are found only by analysis of the data.Crowd sourcing is also needed which is evident by the popularity of apps such as TrueCaller that have been able to build database through effective crowdsourcing. Machine Learning also needs data labelled by humans and crowd sourcing can serve this purpose as well. Systems like Mechanical Turk etc. is evidence of the need for crowd sourcing.

Q. 28. How the cases of false complaints can be mitigated or eliminated? Whether complaints in cases when complainant is in business or commercial relationship with party against which complaint is being made or in case of family or friends may not be entertained? Whether there should be provision to issue notice before taking action and provision to put connection in suspend mode or to put capping on messages or calls till investigation is completed? Please give your suggestions with reasons

If there is a tagged dataset of false complaints along with some meta data and CDRs then a false complaint system can be designed that makes use of an AI algorithm. Since a redressal system has been operational for some time it would be safe to assume that such a tagged dataset is available.The key here again is access to data.TSP can maintain a centralized database of such cases with meta data and "features" that don't reveal PII or if there is enough data at the TSP they can build/outsource building of a Machine Learning algorithm on their data island.Additionally when a resolution happens the database should get updated automatically as part of the resolution so the algorithm can learn automatically and improve over time.

Q. 29. How the scoring system may be developed for UCC on the basis of various parameters using signature solutions of access providers? What other parameters can be considered to detect, investigate and mitigate the sources of UCC? How different access providers can collaborate? Please give your suggestions with reasons.

The scoring system should be a 2-layered system. There should be a fast rule evaluation system that incorpates the output of the signature system if possible along with an ML system as well. The rule system can incorporate domain knowledge and can serve as a first level check before sending data to the Machine Learning algorithm for scoring.Macro/aggregate level data such as avg inter SMS send rate,call talk time ,inferred location,metrics from the call graphetc. of originators along with granular data such as message content, spectral call features etc. can be used to build a Machine Learning algorithm.There is no way of knowing upfront what the Machine Learning algorithm might learn as important parameters so it's useful to send as much data as possible to the Machine Learning algorithm. It is not unusual to have 3000-10000 parameters being looked at in advanced Machine Learning systems.