

EMERGENCY SERVICE – SPRACHBASIERTE KLASSIFIKATION EINGEHENDER ANRUFEN IN AUSNAHMESITUATIONEN

Marcus Petersen¹, Karl-Heinz Niedrist¹, Matthias Busch¹, Florian Marquardt¹, Ingo Siegert²

¹regiocom SE,

²Mobile Dialogsysteme, Institut für Informations- und Kommunikationstechnik,
Otto-von-Guericke-Universität Magdeburg
marcus.petersen@regiocom.com

Kurzfassung: In diesem Beitrag präsentieren wir einen sprachbasierten Ansatz zur automatisierten Klassifikation und Selektion von eingehenden Anrufen einer Kundenhotline in Ausnahmesituationen. Eine Ausnahmesituation tritt z.B. genau dann ein, wenn der Strom in einem Netzgebiet großflächig ausfällt. Davon sind in der Regel viele Unternehmen/Haushalte betroffen, die sich in Folge dessen vermehrt bei der Kundenhotline des zugehörigen Netzbetreibers melden. Woraufhin es dort zu einer Überlastsituation kommt, in der nur noch ein Bruchteil der Anrufe bearbeitet werden kann. Die Auswahl der Anrufe erfolgt dabei zufällig, sodass auch Anrufe von besonderer Wichtigkeit – sog. *Emergency Calls (EC)* – unbearbeitet bleiben, was dramatische Folgen haben kann. In dem hier vorgestellten Lösungsansatz werden daher Sprachinhalt und -prosodie eines jeden Anrufenden analysiert und für die Klassifikation miteinander kombiniert. Das Ergebnis ist eine Klasse, die die Priorität des Anrufs widerspiegelt und als Selektionskriterium genutzt werden kann. Ziel ist es alle *Emergency Calls* zu erkennen, zu selektieren und unmittelbar zu bearbeiten.

1 Ausgangssituation und Handlungsbedarf

Gemäß dem BUNDESMINISTERIUM DES INNEREN gilt eine Infrastruktur als kritisch, wenn diese essentiell für die Gesellschaft ist und „*bei deren Ausfall oder Beeinträchtigung nachhaltig wirkende Versorgungsengpässe, erhebliche Störungen der öffentlichen Sicherheit oder andere dramatische Folgen eintreten würden*“ [1]. Teil der Infrastruktur ist die Stromversorgung, die aufgrund ihrer grundlegenden Stellung im Gesamtsystem zu den sog. technischen Basisinfrastrukturen gehört. Ein länger andauernder Massenausfall dieser Organisationen und Einrichtungen hat besonders gravierende Konsequenzen zur Folge [2], [3].

Kommt es bspw. durch eine Naturkatastrophe (Sturmflut, Herbststurm etc.) zu einem großflächigen Ausfall der Stromversorgung in einem Netzgebiet, so sind in der Regel viele Unternehmen und Privathaushalte davon betroffen, die alle ein berechtigtes Interesse an aktuellen Informationen zu der Störung haben. Zu den gängigen Informationsquellen zählen neben Nachrichtensendungen, Tageszeitungen und Internetseiten insbesondere auch die Kundenhotlines der betroffenen Netzbetreiber.

Naturkatastrophen treten spontan und ungeplant auf, sodass eine kurzfristige Aufstockung der Mitarbeiterinnen und Mitarbeiter an den Hotlines in der Regel nicht möglich ist. Erhöhte Anrufvolumina und lange Wartezeiten sind die Folge. In dieser plötzlich-auf tretenden Überlastsituation können nicht alle Anrufenden bedient werden. Davon sind nicht nur Endkundinnen und Endkunden mit reinem Informationsbedarf („*Wann geht der Kühlschrank wieder, meine Lebensmittel verderben sonst?*“) betroffen, sondern auch Rettungskräfte und Großbetriebe, die schnell aktuelle Informationen benötigen, um Personen- und oder erhebliche Sachschäden abwenden zu können. Für Rettungskräfte ist es bspw. enorm wichtig, Stromabschaltungen veranlassen zu können, Großbetriebe sind wegen der Technik oftmals auf eine kontinuierliche Stromversorgung angewiesen: Wenn z.B. ein Großbauer seine Kühe aufgrund eines länger

andauernden Stromausfalls nicht regelmäßig melken kann, so kann dies für die Kühe schnell lebensbedrohlich werden.

Trotz eines großflächigen Stromausfalls ist in Krisensituationen mit einem erhöhten Anrufaufkommen zu rechnen: So sind Anrufe (engl. *Calls*) aus dem Festnetz mit Telefonen ohne eigenen Stromanschluss problemlos möglich, da diese den benötigten Strom über das Netz beziehen und es dafür eine Notstromversorgung gibt. Für einen begrenzten Zeitraum werden zudem alle relevanten Sende- und Empfangseinheiten des Mobilfunknetzes über Notstrom aufrechterhalten [4].

Während der Notstromversorgung des Mobilfunknetzes kann sich das Anrufaufkommen noch zusätzlich verstärken. Wenn eine Störung nicht das komplette Netzgebiet eines Betreibers abdeckt, dann melden sich auch Endkundinnen und -kunden aus den nicht betroffenen Gebieten, was das Anrufvolumen noch weiter steigert.

Als Sturmtief „Herwart“ am 29. Oktober 2017 großflächig Schäden in Deutschland anrichtete, stieg das Anrufvolumen bei der von der regiocom SE betreuten Hotline um etwa das 8,5-fache des üblichen Werts auf knapp 3.300 Anrufe an. Davon konnten ca. 28% zufällig ausgewählte Anrufe durch eine Mitarbeiterin resp. einen Mitarbeiter entgegengenommen und bearbeitet werden. Hierbei wird genau dann ein Anruf angenommen, wenn gerade freie Kapazitäten im Contact-Center zur Verfügung stehen, eine Priorisierung der Anrufe erfolgt nicht. Durch die spontan auftretenden Naturereignisse ist es nicht möglich eine ausreichende Anzahl an Mitarbeiterinnen resp. Mitarbeitern zu aktivieren, und ein ständiges Vorhalten dieser Notfallkapazitäten ist wiederum nicht betriebswirtschaftlich.

Vor diesem Hintergrund gilt es alle wichtigen Anrufe in der Hotline zu erkennen und priorisiert zu bearbeiten. Klassische Maßnahmen wie eine eigene Rufnummer für diese Kundengruppe oder Telefonautomaten zur Kategorisierung der Kundenanliegen blieben in der Vergangenheit erfolglos, da in der akuten Stresssituation entweder die Nummer vergessen wird oder sich die Anrufenden solange durch das Menü des Automaten probieren, bis sie schließlich mit einem Menschen verbunden werden. Des Weiteren fehlt im Moment eine Möglichkeit, um die Rate der abgelehnten wichtigen Anrufe bestimmen zu können.

Aus der geschilderten Gesamtsituation resultiert folgender Handlungsbedarf:

1. Automatisierte Klassifikation der Anrufenden in *wichtige Anrufe mit Gefahr für Leib und Leben und/oder mit erheblichen Sachschäden* – sog. *Emergency Calls (EC)* und *Anrufe von Informationsbedürftigen (INF)*.
2. Weiterleitung der Anrufenden aus der Klasse *EC* zu den Mitarbeiterinnen und Mitarbeitern der Hotline.
3. Gezielte Informationsbereitstellung für Anrufende der Klasse *INF* über automatisierte Bandansagen zum aktuellen Bearbeitungsstand der Störung.
4. Integration der Lösung in die bestehende Infrastruktur gemäß dem *Gatekeeper-Pattern* (vgl. dazu [5]) zur Sicherung der Erreichbarkeit für *Emergency Calls*.

2 Lösungsansatz und Hypothesen

Im Rahmen dieser Arbeit stellen wir einen sprachbasierten Lösungsansatz vor, der Anrufe in Ausnahmesituationen automatisiert anhand von Kundenanliegen/Inhalt und Prosodie klassifizieren soll (*EC* oder *INF*) und entsprechende Selektionsmaßnahmen vornimmt. Der Ansatz basiert auf der Grundannahme, dass Rettungskräfte ihr Anliegen aufgrund ihrer Erfahrung ruhiger und sachlicher vortragen als Privatpersonen sowie andere Begriffe und Beschreibungen nutzen. Für die Klassifikation soll daher ein automatisierter Sprachinhalts- und Sprachprosodie-Erkennen eingesetzt werden. Ziel ist es alle *Emergency Calls* zu erkennen und un-

mittelbar zu bearbeiten. Wenn kein *Emergency Call* vorliegt, so gilt es selbstredend so viele *Informationsbedürftige* wie möglich zu bedienen.

Die zu entwickelnde Lösung soll unmittelbar nach der Begrüßung in die Kundenhotline integriert werden. Jeder Anrufende wird dann aufgefordert sein Anliegen kurz vorzutragen. Danach erfolgt die Klassifikation, die nahezu in Echtzeit ausgeführt werden muss, um technisch bedingte Wartezeiten zu vermeiden. Als Ergebnis wird dem Anrufenden dann entweder eine automatisierte Bandansage mit aktuellen Informationen vorgespielt oder er wird mit einer Mitarbeiterin resp. einem Mitarbeiter verbunden. Das Ziel ist hier eine minimale Fehlerquote (engl. *False-rejects*) bei der Erkennung von *Emergency Calls* zu erreichen. Durch die Möglichkeit Bandansagen für *Informationsbedürftige* schalten zu können, sind fälschlicherweise durchgeroutete Anrufe (engl. *False-accepts*) tolerierbar, da diese von einer Mitarbeiterin resp. einem Mitarbeiter sehr schnell an die Bandansage weitergeleitet werden können und somit nur sehr kurz die Leitung belegen. Dieser Ansatz gewährleistet das Maximum an Anrufen bei einem spontan-auftretenden Störereignis zu bearbeiten, unter der Prämisse, alle *EC*-Anrufe priorisiert zu bedienen. Abbildung 1 fasst den Lösungsansatz („Solution“) an dem Beispielszenario aus Abschnitt 1 noch einmal kurz zusammen.

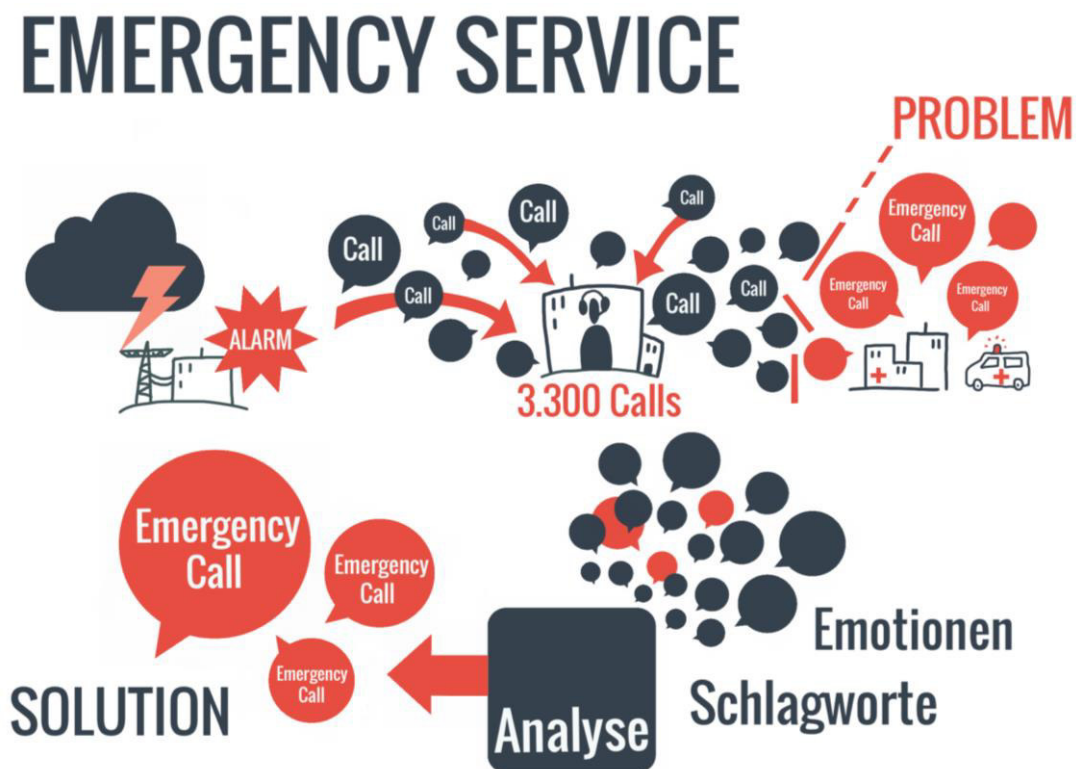


Abbildung 1 – Schematische Darstellung des Lösungsansatzes (Design: Sandra Dengler)

Zusammenfassend gilt es folgende Hypothesen im Rahmen der Entwicklung zu beweisen:

1. *Personen, die regelmäßig mit Ausnahmesituationen zu tun haben, tragen ihr Anliegen ruhiger und gelassener vor als Personen, die derartige Situationen seltener erleben. Zudem bedienen sich diese Personen eines anderen Vokabulars.*
2. *Die automatisierte Klassifikation der Anrufenden in die vorgestellten Klassen (EC und INF) erfordert zwingend eine Berücksichtigung von Inhalt und Prosodie.*

3 Sprachportal „Emergency Service“

Das Sprachportal „Emergency Service“ dient der automatisierten Klassifikation und Selektion eingehender Anrufe in Ausnahmesituationen. Die Architektur des Sprachportals wird in Ab-

schnitt 3.1 skizziert, danach folgen Erläuterungen zu den Besonderheiten bei der Trainings- und Evaluationsdatenerstellung (Abs. 3.2) sowie der Einbettung und Nutzung der Lösung im Rahmen der bestehenden Kundenhotline (Abs. 3.3).

3.1 Lösungsarchitektur

Das Sprachportal „Emergency Service“ besteht aus den beiden Hauptkomponenten *Sprachanalyse* und *Klassifikation* und soll mithilfe der Audiostreams der Anrufenden in die bestehende Kundenhotline resp. Telefonanlage eingebunden werden. So soll das Automatic Call Distribution (ACD)-System die eingehenden Anrufe entgegennehmen, per Bandansage begrüßen und anschließend direkt an das Sprachportal weiterleiten. Das Ergebnis des Portals dient der ACD dann als Routingkriterium: Erkannte *Emergency Calls* sollen direkt auf die zur Verfügung stehenden Mitarbeiterinnen und Mitarbeiter verteilt werden, wohingegen den *Informationsbedürftigen* Störungsinformationen per Bandansage vorgespielt werden sollen. Durch den Einsatz eines Text-to-Speech-Dienstes ist eine automatisierte Erstellung und Aktualisierung der Ansagetexte möglich, sodass die Anrufenden stets aktuelle und ggf. sogar Anliegen-spezifische Informationen erhalten. Die einzelnen Bestandteile des Sprachportals und deren Zusammenspiel im Kontext der *Emergency Call (EC)*-Erkennung zeigt Abbildung 2.

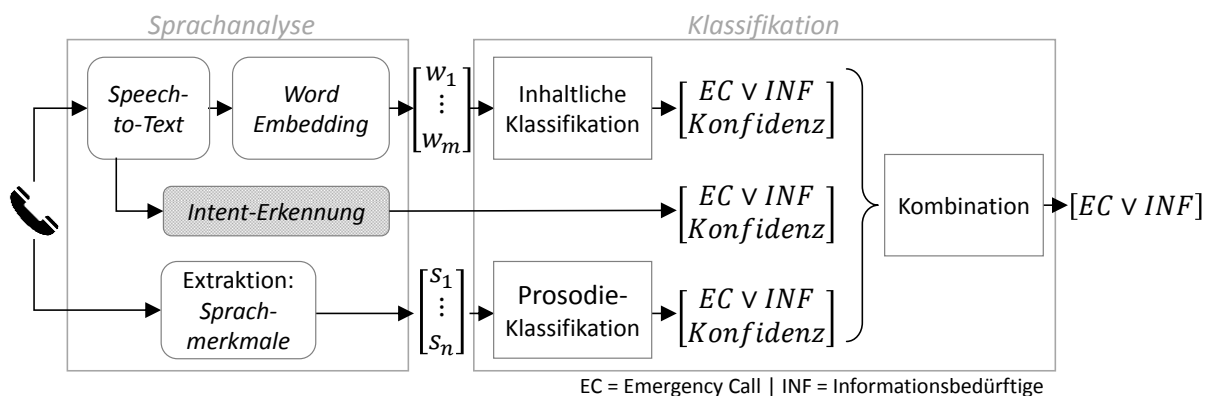


Abbildung 2 – Bestandteile des Sprachportals und deren Zusammenspiel im Kontext der *Emergency Call (EC)*-Erkennung

Die *Sprachanalyse* ist für das Erkennen und Verstehen gesprochener Sprache zuständig. Im ersten Schritt wird die Sprachnachricht des Anrufenden mithilfe eines *Speech-to-Text*-Dienstes in ein Textfragment umgewandelt. Danach wird die Sprachinhaltsanalyse, bestehend aus einer Sprachmodellierung bei der gängige Wörter oder Phrasen von *Emergency Calls* im Sprachinhalt detektiert werden (Häufigkeit, umgebende Wörter) sowie einer Intent-Erkennung („Absichtserkennung“) auf Basis hinterlegter Schlagworte/Trainingsdaten, durchgeführt. Die Modellierung erfolgt mithilfe dynamischer Methoden, wie bspw. *Bag of Words* oder *Word Embedding* [6]. Die Inhalte der *Emergency Calls* werden von den Mitarbeiterinnen und Mitarbeitern des Contact-Centers ermittelt und bereitgestellt. Für die Intent-Erkennung wird auf etablierte Cloud-Dienste, wie bspw. *Google Dialogflow* [7], zurückgegriffen. Diese liefern neben der erkannten Klasse auch den zugehörigen Konfidenzwert, somit liegen alle Informationen für die abschließende Kombination der Einzelergebnisse vor. Des Weiteren erfolgt im Rahmen der *Sprachanalyse* auch die Auswertung der Sprachprosodie. Im Gegensatz zur Intent-Erkennung wird hier lokal eine Merkmalsextraktion durchgeführt, wobei auf etablierte Merkmale zum Erkennen von Sprachemotionen zurückgegriffen wird [8]. Im vorliegenden Fall geht es – anders als bei klassischen Ansätzen – nicht darum eine bestimmte Emotionsklasse zu detektieren [9], verschiedene Emotionsklassen voneinander zu unterscheiden [10] oder Unterschiede in der Emotionalität der Dialogpartnerinnen resp. -partner zu erkennen [11, 12], sondern das „Fehlen“ von Emotionalität gegenüber dem Vorhandensein von Emotion

(hohe Erregung, negative Valenz) zu detektieren. Der Vorteil hierbei ist, dass auch für wenig ausgeprägte Emotionen, wie dies in natürlicher Sprache zu erwarten ist [13], eine Unterscheidung von der neutralen Ausgangssituation möglich ist. Als Ergebnis liefert die Komponente alle relevanten Parameter für die Klassifikation. Dies sind zum einen die beiden sog. *Merkmalsvektoren* (engl. *Feature sets*) mit den Wort- (w) und Sprachmerkmalen (s) und zum anderen ist das der Intent des Anrufenden, falls dieser korrekt erkannt werden konnte.

Im Rahmen der **Klassifikation** werden zunächst die Merkmale der Vektoren gemäß der gewählten Klassifikationsverfahren verarbeitet. Durch die Problemstellung, dass große Mengen Anrufe anfallen, ergibt sich für die Klassifikation die Chance, auf Methoden aus dem Machine Learning/Deep-Learning zu setzen. Das Ergebnis besteht jeweils aus der erkannten Klasse (*EC* oder *INF*) und dem zugehörigen Konfidenzwert. Die endgültige Klassifikation des Anrufes erfolgt anschließend durch die gewichtete Kombination der drei Klassifizierungsergebnisse (Sprachmodellierung, Intent und Prosodie), wobei die jeweilige Konfidenz in die Gewichtung einfließt. Damit ist es möglich, Anrufe ohne aussagekräftige Prosodie-Erkennung immer noch anhand des Sprachinhalts einer der beiden Klassen zuzuordnen. Die resultierende Klasse dient der ACD als Routingkriterium.

3.2 Erstellen und Labeln von Trainings- und Evaluationsdatensätzen

Ein wesentlicher Aspekt bei der Erstellung automatisierter Erkennungssysteme ist das Bereitstellen von Daten mit entsprechender Grundwahrheit – den *Labels*. Hierbei muss darauf Wert gelegt werden, dass die Qualität der Labels ausreichend hoch ist [14]. Gerade bei der Verwendung von echten Gesprächsaufzeichnungen einer Kundenhotline ist es von enormer Wichtigkeit, dass die personenbezogenen Daten konsequent geschützt werden. Zu den personenbezogenen Daten gehören zum einen alle Informationen, die zur direkten Unterscheidung oder Rückverfolgung der Identität einer Person verwendet werden können, wie z.B. Name, Sozialversicherungsnummer, Geburtsdatum und -ort oder biometrische Aufzeichnungen; und zum anderen alle anderen Informationen, die mit einer Person indirekt verbunden oder verknüpfbar sind, wie z.B. medizinische, erzieherische, finanzielle und/oder Beschäftigungsinformationen [15].

Vor diesem Hintergrund ist eine Bearbeitung der Anrufrdaten außerhalb der regiocom SE nur nach vorangegangener Anonymisierung möglich. Der Prozess ist allerdings sehr zeitaufwendig, da alle Anrufe komplett manuell kontrolliert werden müssen. Alternativ muss die Bearbeitung intern erfolgen. Für die Erstellung der Label ist dies der favorisierte Weg, da es mit den Qualitätsbeauftragten (QB) im Rahmen der Hotline bereits ausgewählte Mitarbeiterinnen und Mitarbeiter gibt, zu deren täglichen Aufgaben das Analysieren von Gesprächsaufzeichnungen zu Qualitätssicherungs- und Trainingszwecken gehört. Daher sind diese Mitarbeiterinnen und Mitarbeiter prädestiniert für das Erstellen der Grundwahrheit der Aufzeichnungen – auch ohne vorherige Anonymisierung.

3.3 Selektion und Routing der eingehenden Anrufe

Die Selektion der Anrufe, die zu einer Mitarbeiterin resp. einem Mitarbeiter durchgestellt werden, ist im ersten Schritt trivial, da zunächst immer versucht wird, alle *Emergency Calls* zu vermitteln. Hierzu wird auf die bestehende Applikationslandschaft zurückgegriffen. Diese stellt den aktuellen Status jeder einzelnen Person („Am Telefon“, „In der Schriftbearbeitung“ etc.) in Echtzeit dar, wobei diese Informationen bereits heute zum Routing innerhalb der ACD genutzt werden. Zudem ermöglicht die Technik eine schnelle und sichere Weitergabe der Anrufe per *Session Initiation Protocol (SIP)* [16].

Interessant wird der Routingansatz erst dann, wenn mehr Anrufe der Klasse *EC* vorliegen, als freie Mitarbeiterinnen und Mitarbeiter zur Verfügung stehen. Für diesen Fall gibt es, naiv be-

trachtet, zwei Lösungsstrategien: So kann, wie heutzutage bei Kundenhotlines üblich, ein Wartefeld für *Emergency Calls* aufgebaut werden, welches als Puffer dient. Allerdings würde dieser Ansatz nicht zu den Anforderungen passen, da dadurch Zeitverzögerungen bei der Bearbeitung der Anrufe auftreten können und diese völlig unabhängig von den konkreten Kundenanliegen wären. Eine Alternative ist die Einführung einer Abstufung / Priorisierung zwischen den aktuell im System vorliegenden *Emergency Calls*. Dabei wäre eine unmittelbare Weiterleitung – *Fastlane* – für Anrufe, die mit einer sehr hohen Konfidenz als *EC* klassifiziert wurden und bei denen außerdem der Inhalt auf „*Gefahr im Verzug schließen lässt*“, denkbar.

4 Evaluation

Die Evaluation umfasst zum einen das Überprüfen der trainierten Modelle und zum anderen das Untersuchen und Bewerten der Hypothesen. Für das Evaluieren der Modelle kann auf Standardmethoden, wie bspw. die 10-Fold-Cross-Validation, zurückgegriffen werden.

Die erste Hypothese „*Personen, die regelmäßig mit Ausnahmesituationen zu tun haben, tragen ihr Anliegen ruhiger und gelassener vor*“ soll mithilfe der Mitarbeiterinnen und Mitarbeiter der Kundenhotline sowie den zugehörigen Qualitätsbeauftragten (QB) als Qualitätssicherungsinstanz belegt werden. Dazu wird eine Feedback-Schleife in Anlehnung an überwachte Online-Lernverfahren (vgl. [17]) implementiert, die es den Mitarbeiterinnen und Mitarbeitern über ein Auswahlmenü erlaubt, jeden Anruf nach erfolgreicher Bearbeitung händisch neu zu klassifizieren (falls erforderlich). Diese manuelle Zuordnung muss durch einen QB bestätigt werden, bevor die Klassifizierungskomponente entsprechend neu trainiert wird. Sollte sich dieses Vorgehen erfolgreich etablieren, so ist eine kontinuierliche Anpassung des Sprachportals an aktuelle Begebenheiten möglich.

Für die Annahme das „*die automatisierte Klassifikation der Anrufenden in die vorgestellten Klassen [...] zwingend eine Berücksichtigung von Inhalt und Prosodie [erfordert]*“ sollen zwei unterschiedliche Klassifizierungskomponenten trainiert werden – mit und ohne Prosodie-Erkennung. Anschließend werden beide Komponenten mit derselben Evaluationsmenge konfrontiert und das Ergebnis miteinander verglichen.

5 Zusammenfassung und Ausblick

Tritt ein großflächiger Stromausfall in einem Netzgebiet auf, so kommt es bei den betroffenen Netzbetreibern zwangsläufig zu Überlastungen der Kundenhotlines. In der Konsequenz können nicht alle Anrufe bearbeitet werden, was insb. für nicht angenommene *Emergency Calls* kritisch sein kann. Hierfür wurde ein sprachbasierter Lösungsansatz vorgestellt, der die Anrufe auf Basis von Inhalt und Prosodie klassifiziert und entsprechend routet. Dabei wurde insbesondere auf die Themen Klassifikation, Erstellen und Labeln notwendiger Daten sowie Evaluation eingegangen.

Der Lösungsansatz wurde bislang in Teilen umgesetzt. Eine ganzheitliche Validierung steht noch aus. Insbesondere gilt es, die vorgeschlagenen Methoden zur Sprachanalyse und Klassifizierung zu evaluieren und hinsichtlich Praxistauglichkeit zu überprüfen. Eine denkbare Erweiterung des Ansatzes ist die gezielte Abfrage der Postleitzahl (PLZ) des Anrufenden, mit dem Ziel, diese als Identifikator für neue Großstörungen zu nutzen: Wenn während einer großflächigen Störung eine PLZ, die nach aktuellem Kenntnisstand nicht betroffen ist, in einem relativ kurzen Zeitraum häufig auftritt, so ist entweder von einer Ausbreitung der bekannten Störung oder einer neuen Massenstörung auszugehen. Zur Klärung sollte daher mindestens ein Anruf mit der PLZ von einer Mitarbeiterin oder einem Mitarbeiter bearbeitet werden, was eine Adaption der vorgestellten Klassifikation zur Folge hätte. Für die Messung der Häufigkeit gilt es einen geeigneten Schwellwert zu definieren.

Literatur

- [1] BUNDESMINISTERIUM DES INNEREN (BMI): Schutz Kritischer Infrastrukturen – Basischutzkonzept: Empfehlungen für Unternehmen. Bundesministerium des Inneren, Berlin, 2005, S. 51.
- [2] BUNDESMINISTERIUM DES INNEREN (BMI): Nationale Strategie zum Schutz Kritischer Infrastrukturen (KRITIS-Strategie). Bundesministerium des Inneren, Berlin, 2009, S. 5.
- [3] BIRKMANN, J.; BACH, C.; GUHL, S.; WITTING, M.; WELLE, T.; SCHMUDE, M.: State of the Art der Forschung zur Verwundbarkeit kritischer Infrastrukturen am Beispiel Strom, Stromausfall. Schriftenreihe Forschungsforum Öffentliche Sicherheit, Nr. 2, Freie Universität Berlin, Berlin, 2010, S. 13.
- [4] PETERMANN, T.; BRADKE, H.; LÜLLMANN, A.; POETZSCH, M.; RIEHM, U.: Gefährdung und Verletzbarkeit moderner Gesellschaften – am Beispiel eines großräumigen Ausfalls der Stromversorgung. Endbericht zum TA-Projekt, Arbeitsbericht Nr. 141, Büro für Technikfolgen-Abschätzung beim Deutschen Bundestag, 2010, S. 79-89.
- [5] RAJ, P.; RAMAN, A.; SUBRAMANIAN, H.: Architectural Patterns – Uncover essential patterns in the most indispensable realm of enterprise architecture, Packt Publishing Ltd., Birmingham, 2017, S. 371.
- [6] LUQUE, J.; SEGURA, C.; SANCHEZ, A.; UMBERT, M.; GALINDO, L.: The role of linguistic and prosodic cues on the prediction of self-reported satisfaction in contact centre phone calls. In: Proc. of the INTERSPEECH-2017, 2017, S. 2346-2350.
- [7] SINGH, A.; RAMASUBRAMANIAN, K.; SHIVAM, S.: Introduction to Microsoft Bot, RASA, and Google Dialogflow. In Building an Enterprise Chatbot: Work with Protected Enterprise Data Using Open Source Frameworks, Apress, Berkeley 2019.
- [8] VERVERIDIS, D.; KOTROPOULOS, C.: Emotional speech recognition: Resources, features, and methods. Speech Communication, 48 (9), 2016, S. 1162-1181.
- [9] DEVILLERS, L.; VIDRASCU, I.: Real-life emotions detection with lexical and paralinguistic cues on human-human call center dialogs. In Proc. of the INTERSPEECH-2006. Pittsburgh, USA, S. 801–804, 2006.
- [10] SCHULLER, B.; BATLINER, A.; STEIDL, S.; SEPPI, D.: Recognising realistic emotions and affect in speech: State of the art and lessons learnt from the first challenge. Speech Communication 53 (9-10), 2011, S. 1062-1087.
- [11] MOLINO, M.; EMANUEL, F.; ZITO, M.; GHISLIERI, C.; COLOMBO, L.; CORTESE, C.: Inbound Call Centers and Emotional Dissonance in the Job Demands – Resources Model. Frontiers in Psychology, Vol. 7, 2019.
- [12] SIEGERT, I.; PHILIPPOU-HÜBNER, D.; TORNOW, M.; HEINEMANN, R.; WENDEMUTH, A.; OHNEMUS, K.; SCHREIBER, G.: Ein Datenset zur Untersuchung emotionaler Sprache in Kundenbindungsdialogen. In: Tagungsband der 26. Konferenz Elektronische Sprachsignalverarbeitung 2015, S. 180-187.
- [13] SIEGERT, I.; LOTZ, A.; EGOROW, O.; BÖCK, R.; SCHEGA, L.; TORNOW, M.; THIERS, A.; WENDEMUTH, A.: Akustische Marker für eine verbesserte Situations- und Intentionserkennung von technischen Assistenzsystemen. In: Technische Unterstützungssysteme, die die Menschen wirklich wollen: Zweite Transdisziplinäre Konferenz, 2016, S. 465-474.
- [14] SIEGERT, I.; BÖCK, R.; WENDEMUTH, A.: Inter-rater reliability for emotion annotation in human-computer interaction - comparison and methodological improvements. Journal on multimodal user interfaces, 8 (1), 2013, S. 17-28.
- [15] Art. 4 Abs. 1, Verordnung 2016/679 (EU), Datenschutz-Grundverordnung.
- [16] RUSSELL, T.: Session Initiation Protocol (SIP): Controlling Convergent Networks. 1st ed., McGraw-Hill Osborne Media, 2008.

- [17] ŽLIOBAITĖ I.; BIFET A.; PFAHRINGER B.; HOLMES G.: Active Learning with Evolving Streaming Data. In: Gunopulos D., Hofmann T., Malerba D., Vazirgiannis M. (eds) Machine Learning and Knowledge Discovery in Databases. ECML PKDD 2011, Lecture Notes in Computer Science, Vol. 6913, Springer, Berlin, Heidelberg, 2011.