1 Aufgabe des Programmes

Das im Praktikum zu erstellende Programm hat die Aufgabe, aus den Wortlisten zweier beteiligter Parteien einen Klassifikator für Spammails zu erzeugen, den jede Partei einzeln verwenden kann.
Es gibt bei der Verwendung des Programmes genau zwei beteiligte Parteien. Diese werden im folgenden als Anwender bezeichnet. Jeder dieser beiden Anwender muss fuer das Programm als Eingabe eine Liste von Worten und eine Menge von E-Mail Inhalten bereitstellen. Diese Wortlisten enthalten einzelne Worte, welche der entsprechende Anwender in der Spamklassifikation verwendet sehen will. Die Menge von Email Inhalten enthält Inhalte von E-Mails, die gesammelt und vorklassifiziert wurden. Der berechnete Klassifikator ist in der Lage, Mails in Spam und kein Spam einzuteilen und kann von jedem Anwender ohne die anderen Anwender verwendet werden.
Dadurch ist es beispielsweise möglich, dass Betreiber von E-Mail-Servern gemeinsam bessere E-Mail-Klassifikatoren berechnen, als jeder einzelne mit seinen vorklassifizierten E-Mails es koennte. Danach kann von jedem Anwender der berechnete Klassifikator verwendet werden, um den weiteren E-Mailverkehr zu klassifizieren.

2 Garantierter Level der Geheimhaltung

Das Programm garantiert in keiner Weise für die Geheimhaltung der Wortliste. Es ist jedoch dem Anwender bekannt, dass die Wortliste publiziert werden wird, sodass es jedem Anwender möglich ist, keine geheimzuhaltenden Informationen in die Wortliste zu übernehmen.
Das Programm garantiert nur dafür, dass ueber die Menge der vorklassifizierten E-Mail-Inhalte nur soviele Informationen preisgegeben werden, wie aus dem resultierenden Klassifikator entnehmbar sind. Da wir den ID3-Algorithmus verwenden, ist dies umformulierbar zu: Es wird potentiell für alle Worte der Wortliste eines einzelnen Anwenders preisgegeben, wie häufig dieses Wort in der Vereinigung beider Mengen von E-Mail Inhalten vorkommt. Da jedoch jeder Benutzer selbst bestimmen kann, welche Worte in der Wortliste auftauchen, wurde dieser Informationsverlust als akzeptabel eingestuft, da er unvermeidbar ist.
Die Publizierung der Wortliste ist unvermeidbar, da beide Parteien am Ende einen Klassifikator erhalten, den sie alleine benutzen. Das bedeutet, beide Parteien haben am Ende vollständige Informationen über den Klassifikator. In unserem Fall wird ein Entscheidungsbaum berechnet. Das bedeutet insbesondere, dass die Attribute des Baumes jedem Anwender vollständig bekannt sind. Da die Attribute des Baumes aus der Wortliste jedes Anwenders berechnet werden, muss somit damit gerechnet werden, dass die vollständige Wortliste der Benutzer preisgegeben wird.
Der Informationsverlust durch den Klassifikator kann ebenfalls nicht weiter reduziert werden. Es ist so, dass in der Lernphase eines Klassifikators Informa-

tionen aus den Trainingsdaten des Klassifikators in den Klassifikator einkodiert werden. Dies ist notwendig, damit der Klassifikator im folgenden ohne die Trainingsmenge weiterhin klassifizieren kann. Damit und aus dem Fakt, dass jeder Anwender den Klassifikator alleine vollständig verwenden können muss, folgt, dass jeder Informationen über die Gesamtmenge aller E-Mails enthalten wird, da dies durch die Natur eines Klassifikators und die vollständigen Informationen über den Klassifikator impliziert wird.

3 Sicherheit im $ID3_\delta$

Die Sicherheit aller Bestandteile des Algorithmus' in dem Sinne, dass man über die Datenbank des anderen Anwenders nicht mehr erfährt als der Baum preisgibt, ist im Paper bewiesen. Damit bleibt nur noch zu klären wie die Attribute gefunden werden.

4 Finden der Attribute des Baums

Es gab 3 Ansätze zum Finden der Attribute:

- alle Worte aus den E-mail-Inhalten werden genutzt
- das gesamte Wörterbuch wird genutzt
- beide Partner geben eine Wortliste vor

Bei Option 1 ist das Problem, dass der andere Anwender eine Vorstellung davon gewinnt, welche Worte in den Nutzmails oft vorkommen, sodass die Geheimhaltung der Inhalte der E-Mails stärker gefährdet ist.
Option 2 hat ein ähnliches Problem: Alle Worte im Lexikon sind mit grosser Wahrscheinlichkeit eine Obermenge aller Worte, die in den E-Mails vorkommen. Damit ist analog zu Option 3 die Geheimhaltung der Inhalte der E-Mails stärker gefährdet.
Bei Option 3 wird ebenfalls die Geheimhaltung der Inhalte der E-Mails gefährdet, jedoch hat jeder Anwender hier die direkte und uneingeschränkte Kontrolle über diesen Verlust der Geheimhaltung.
Einen Informationsgewinn des anderen Anwenders kann man nicht verhindern, da dieser alle Worte kennen muss um das nächste Attribut für den Baum zu bestimmen und durch das Lernen des Klassifikators auf jeden Fall genaues Wissen darüber erhält, welche Attribute die Mails gut klassifizieren.
Generell gilt, dass die Wahl eines Attributes Informationen an den anderen Anwender übertrgt: mit hoher Wahrscheinlichkeit ist dieses Attribut, welches gleichbedeutend mit einem Wort ist, für den anderen (auch) wichtig. Also können wir nur vermeiden, dass ungewollt Informationen weitergegeben werden.
Wir definieren privat in dem Sinne dass jeder Anwender selbst entscheiden kann welche Informationen weitergegeben werden. Das fordert — verglichen mit dem Paper — weniger, sodass die Sicherheit des $Id3_\delta$ nicht neu bewiesen werden

muss.
Option 3 erfüllt diese Bedingung; wir gehen im Folgenden davon aus dass beide eine Wortliste aufstellen.

5 Finden der gemeinsamen Wortliste

Es bleibt zu klären, wie die Wortlisten zusammengeführt werden. Dann sollte die preisgegebene Information über die Wortliste minimal sein, auch wenn im Folgenden demonstriert werden wird, dass keine Geheimhaltung der Wortliste garantiert werden kann.
Es seien oBdA. zwei Anwender beteiligt und W_1 die Wortliste des ersten Anwenders und W_2 die Wortliste des zweiten Anwenders und *join* sei eine Operation die die beiden Wortlisten zusammenführt. Dann sollten sowohl $join(W_1, W_2) \subseteq W_1$ als auch $join(W_1, W_2) \subseteq W_2$ gelten. Damit ergibt sich der Schnitt der Wortlisten als einfache Implementierung von *join*.