Description of system performance

AC

Contents

A first attempt 1
Read in data 1
Explaining misses and false alarms across files based on file characteristics (Marvin VAD on BabyTrain - 5 class, old architecture) 4

A first attempt

The goal here is to link system performance to characteristics of files and/or speakers on files. See this file for explanation of the fields.

Read in data

We read in background data (only exists for babytrain), data describing files and speakers, as well as results. For now, we only have results at the level of files (not individual speakers within files).

```
read.csv("../BabyTrain_ages.csv")->ages
ages[ages$corpus!="corpus",]->ages
ages$age=as.numeric(as.character(ages$age))

allres=dir("../extraction/results/")

datsp=NULL
for(j in allres[grep("perSpeaker",allres)])  datsp=rbind(datsp,cbind(j,read.csv(paste0("../extraction/r

datf=NULL
for(j in allres[grep("perSpeaker",allres,invert=T)])  datf=rbind(datf,cbind(j,read.csv(paste0("../extra

merge(datsp,ages,by.x="file",by.y="basename",all.x=T)->datsp
merge(datf,ages,by.x="file",by.y="basename",all.x=T)->datf

summary(datsp)
```

```
##                          file                                    j
##  namibia_aiku_20161111_19980:   9   BabyTrain_train_perSpeaker.csv:4163
##  namibia_eiun_20161113_30780:   9   BabyTrain_dev_perSpeaker.csv  :2071
##  namibia_oegd_20161109_23580:   8   BabyTrain_test_perSpeaker.csv :1254
##  namibia_oekd_20160712_12780:   8   AMI_train_perSpeaker.csv      : 471
##  namibia_oekd_20160712_16380:   8   AMI_dev_perSpeaker.csv        : 104
##  namibia_oekd_20170308_19980:   8   AMI_test_perSpeaker.csv       :  95
##  (Other)                    :8188   (Other)                       :  80
##     speaker         role      tot_ovl_speech      tot_nonovl_speech
##  MOT*   : 925   FEM   :2617   Min.   :   0.0000   Min.   :-297.429
##  C1     : 909   MAL   :1370   1st Qu.:   0.0000   1st Qu.:   0.000
##  MA1    : 563   CHI   :1375   Median :   0.0308   Median :   0.000
##  FA1    : 551   KCHI  :2329   Mean   :  33.5954   Mean   :  39.993
```

```
##  C2     : 414   SPEECH: 547   3rd Qu.:   2.9493   3rd Qu.:   7.659
##  UU     : 378                 Max.   :2688.8500   Max.   :2665.000
##  (Other):4498
##    snr             corpus              child.id         age
##  []  :8174   namibia  :3977   fhugo        : 208   Min.   : 1.00
##  NA's:  64   lena_lyon: 907   nath         : 122   1st Qu.:13.00
##              tsay     : 751   marin        : 118   Median :27.00
##              paido    : 649   uebn_20170309: 115   Mean   :25.04
##              tsimane  : 447   ern          : 109   3rd Qu.:33.00
##              (Other)  : 330   (Other)      :6389   Max.   :66.00
##              NA's     :1177   NA's         :1177   NA's   :1177
```

```
summary(datf)
```

```
##                  file                         j        key_child_age
##  ES2003a.Mix-Headset:   1   BabyTrain_train.csv:1544   Mode:logical
##  ES2003b.Mix-Headset:   1   BabyTrain_dev.csv  : 736   NA's:2881
##  ES2003c.Mix-Headset:   1   BabyTrain_test.csv : 413
##  ES2003d.Mix-Headset:   1   AMI_train.csv      : 118
##  ES2011a.Mix-Headset:   1   AMI_dev.csv        :  26
##  ES2011b.Mix-Headset:   1   AMI_test.csv       :  24
##  (Other)            :2875   (Other)            :  20
##   clip_length      nb_diff_speakers  nb_children      nb_fem_ad
##  Min.   :   60.0   Min.   :1.000    Min.   :0.000   Min.   :0.0000
##  1st Qu.:   60.0   1st Qu.:1.000    1st Qu.:1.000   1st Qu.:0.0000
##  Median :   60.0   Median :3.000    Median :1.000   Median :1.0000
##  Mean   :  467.6   Mean   :2.859    Mean   :1.286   Mean   :0.9084
##  3rd Qu.:  300.0   3rd Qu.:4.000    3rd Qu.:2.000   3rd Qu.:1.0000
##  Max.   :10723.0   Max.   :9.000    Max.   :4.000   Max.   :4.0000
##
##    nb_mal_ad       nb_uncertain    prop_ovl_speech   prop_nonovl_speech
##  Min.   :0.0000   Min.   :0.0000   Min.   :0.00000   Min.   :0.2000
##  1st Qu.:0.0000   1st Qu.:0.0000   1st Qu.:0.00000   1st Qu.:0.9200
##  Median :0.0000   Median :0.0000   Median :0.01000   Median :0.9900
##  Mean   :0.4755   Mean   :0.1899   Mean   :0.06742   Mean   :0.9326
##  3rd Qu.:1.0000   3rd Qu.:0.0000   3rd Qu.:0.08000   3rd Qu.:1.0000
##  Max.   :4.0000   Max.   :1.0000   Max.   :0.80000   Max.   :1.0000
##
##   avg_voc_dur           snr                 corpus        child.id
##  Min.   :   0.42   Min.   :  0.03003   namibia  :1062   fhugo  :  71
##  1st Qu.:  11.87   1st Qu.:  0.71457   paido    : 649   nath   :  48
##  Median :  24.70   Median :  0.84570   lena_lyon: 323   ern    :  36
##  Mean   : 190.21   Mean   : 13.56329   tsay     : 237   flore  :  36
##  3rd Qu.:  61.69   3rd Qu.:  1.96800   tsimane  : 154   leon   :  36
##  Max.   :7659.23   Max.   :104.58705   (Other)  : 122   (Other):2320
##                    NA's   :34          NA's     : 334   NA's   : 334
##       age
##  Min.   : 1.00
##  1st Qu.:15.50
##  Median :30.00
##  Mean   :29.19
##  3rd Qu.:39.00
##  Max.   :66.00
##  NA's   :334
```

```r
#read.table("../AMI_MixHeadset__test__v1.txt", skip=2)->out
read.table("../eval.txt", skip=2)->out
#colnames(out)<-c("file","der","purity","coverage","total","correct","cor.pc","fa","fa.pc","miss","miss

colnames(out)<-c("file","der","accuracy","precision","recall","total","fa","fa.pc","miss","miss.pc")


merge(out,datf,all=T)->out

summary(out)
```

```
##                                 file           der
##  aclew_starter_ROS_9559_15_01_03600:   1   Min.   :   1.54
##  aclew_starter_SOD_9733_11_01_36000:   1   1st Qu.:  14.70
##  aclew_starter_WAR_9755_03_01_30816:   1   Median :  36.01
##  lena_lyon_day1_nohlan_10h35       :   1   Mean   :  49.96
##  lena_lyon_day1_nohlan_10h40       :   1   3rd Qu.:  55.90
##  lena_lyon_day1_nohlan_11h35       :   1   Max.   :1258.16
##  (Other)                           :2876   NA's   :2468
##     accuracy       precision          recall           total
##  Min.   :37.18   Min.   :  0.00   Min.   :  0.00   Min.   :    0.97
##  1st Qu.:77.75   1st Qu.: 76.03   1st Qu.: 66.64   1st Qu.:   19.81
##  Median :86.68   Median : 89.75   Median : 81.72   Median :   37.30
##  Mean   :84.93   Mean   : 82.56   Mean   : 77.59   Mean   :  389.02
##  3rd Qu.:93.78   3rd Qu.: 96.14   3rd Qu.: 94.87   3rd Qu.:   66.45
##  Max.   :99.68   Max.   :100.00   Max.   :100.00   Max.   :80526.23
##  NA's   :2468    NA's   :2468     NA's   :2468     NA's   :2468
##        fa               fa.pc              miss
##  Min.   :   0.000   Min.   :   0.000   Min.   :   0.000
##  1st Qu.:   1.468   1st Qu.:   3.223   1st Qu.:   1.418
##  Median :   3.900   Median :   8.700   Median :   5.060
##  Mean   :  35.521   Mean   :  27.545   Mean   :  37.258
##  3rd Qu.:   9.770   3rd Qu.:  20.962   3rd Qu.:  13.365
##  Max.   :7352.780   Max.   :1174.130   Max.   :7712.340
##  NA's   :2468       NA's   :2468       NA's   :2468
##     miss.pc                          j        key_child_age
##  Min.   :  0.000   BabyTrain_train.csv:1544   Mode:logical
##  1st Qu.:  5.133   BabyTrain_dev.csv  : 736   NA's:2882
##  Median : 18.285   BabyTrain_test.csv : 413
##  Mean   : 22.412   AMI_train.csv      : 118
##  3rd Qu.: 33.355   AMI_dev.csv        :  26
##  Max.   :100.000   (Other)            :  44
##  NA's   :2468      NA's               :   1
##   clip_length      nb_diff_speakers  nb_children      nb_fem_ad
##  Min.   :   60.0   Min.   :1.000    Min.   :0.000   Min.   :0.0000
##  1st Qu.:   60.0   1st Qu.:1.000    1st Qu.:1.000   1st Qu.:0.0000
##  Median :   60.0   Median :3.000    Median :1.000   Median :1.0000
##  Mean   :  467.6   Mean   :2.859    Mean   :1.286   Mean   :0.9084
##  3rd Qu.:  300.0   3rd Qu.:4.000    3rd Qu.:2.000   3rd Qu.:1.0000
##  Max.   :10723.0   Max.   :9.000    Max.   :4.000   Max.   :4.0000
##  NA's   :1         NA's   :1        NA's   :1       NA's   :1
##    nb_mal_ad       nb_uncertain    prop_ovl_speech   prop_nonovl_speech
##  Min.   :0.0000   Min.   :0.0000   Min.   :0.00000   Min.   :0.2000
##  1st Qu.:0.0000   1st Qu.:0.0000   1st Qu.:0.00000   1st Qu.:0.9200
```

```
##  Median :0.0000   Median :0.0000   Median :0.01000   Median :0.9900
##  Mean   :0.4755   Mean   :0.1899   Mean   :0.06742   Mean   :0.9326
##  3rd Qu.:1.0000   3rd Qu.:0.0000   3rd Qu.:0.08000   3rd Qu.:1.0000
##  Max.   :4.0000   Max.   :1.0000   Max.   :0.80000   Max.   :1.0000
##  NA's   :1        NA's   :1        NA's   :1         NA's   :1
##   avg_voc_dur           snr                  corpus        child.id
##  Min.   :   0.42   Min.   :  0.03003   namibia  :1062   fhugo  :  71
##  1st Qu.:  11.87   1st Qu.:  0.71457   paido    : 649   nath   :  48
##  Median :  24.70   Median :  0.84570   lena_lyon: 323   ern    :  36
##  Mean   : 190.21   Mean   : 13.56329   tsay     : 237   flore  :  36
##  3rd Qu.:  61.69   3rd Qu.:  1.96800   tsimane  : 154   leon   :  36
##  Max.   :7659.23   Max.   :104.58705   (Other)  : 122   (Other):2320
##  NA's   :1         NA's   :35          NA's     : 335   NA's   : 335
##       age
##  Min.   : 1.00
##  1st Qu.:15.50
##  Median :30.00
##  Mean   :29.19
##  3rd Qu.:39.00
##  Max.   :66.00
##  NA's   :335
```

The table out has a combination of results and descriptors at the level of files.

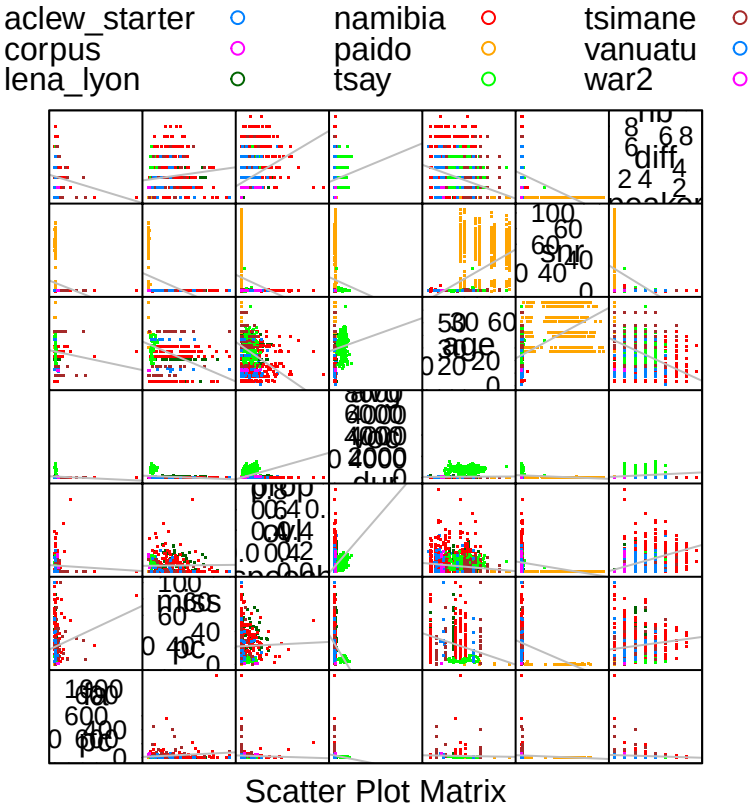
Explaining misses and false alarms across files based on file characteristics (Marvin VAD on BabyTrain - 5 class, old architecture)

A scatter plot matrix shows many bivariate plots. In the one below, we focus exclusively on descriptors at the level of the file and only for BabyTrain because that's what I drew results for. We only have false alarms and misses because we are looking at a VAD system. (In particular, this is Marvin's system for last week.)

```
library(lattice)
```

```
## Warning: package 'lattice' was built under R version 3.4.4
```

```
selected=c("fa.pc","miss.pc","prop_ovl_speech","avg_voc_dur","age","snr","nb_diff_speakers")
selnames=gsub(".","\n",gsub("_","\n",selected),fixed=T)
splom(out[c(selected)],pch=".",groups=out$corpus,varnames=selnames,auto.key = list(columns = 3),axis.li
panel = function(x, y, ...) {
           panel.xyplot(x, y, ...)
          fm <- lm(y ~ x)
          panel.abline(fm,col.line = "gray")
}
)
```



Focus on the last two rows, which show the correlations between percent misses (penultimate row) or percent false alarms (last row) and the following selected characteristics (from left to right):

- proportion of speech that is overlapping
- average vocalization/utterance/sentence duration
- key child age
- SNR calculated as RMS(x_speech)/RMS(x_sil) where x_speech is an array with all the areas of speech in the gold annotation, and x_sil is un array with all the areas of silence
- number of different speakers

So focusing on the last row, false alarms look unrelated to all of these predictors, although this may be because the scale is too large.

One row up, misses does not relate to proportion overlap or number of different speakers, but is anticorrelated with the duration of speech, child age, and SNR.