

Une méthode "inter-subjective" pour l'évaluation de la qualité sonore de voix sur IP

Samir Mohamed¹
Samir.Mohamed@irisa.fr

Francisco Cervantes²
Cervante@itam.mx

Hossam Afifi³
Hossam.Afifi@int-evry.fr

¹ INRIA/IRISA
bureau E318, campus de Beaulieu, 35042 Rennes CEDEX 9, France

² Instituto Tecnológico Autónomo de México (ITAM)
Computer Engineering Department; Rio Hondo No. 1 Mexico D.F., CP-01000 MEXICO

³ Institut National de Télécommunication (INT)
9 rue Charles Fourier, 91011 Evry cedex, France.

Résumé

L'avancée dans le domaine de la transmission et des réseaux soulève un très grand besoin d'estimation de la qualité des flux multimédias en général et audio temps réel en particulier. Parmi les méthodes utilisées, il y a le test subjectif et le test objectif mais il est admis et acquis par la communauté concernée, que seules les méthodes subjectives sont celles qui donnent des résultats réels. C'est d'ailleurs la méthode adoptée dans le test MOS (Mean Opinion Score) utilisé par l'UIT-T dans ses mesures.

Dans cet article, nous proposons une méthode basée sur les réseaux à neurones pour modéliser la façon par laquelle les auditeurs évaluent le son transmis et sujet à des distorsions dues au réseau de transport (ex: taux de pertes de paquets, modèle des pertes, et tailles des paquets). La famille de l'algorithme fait aussi partie des considérations dans notre méthode. Notre algorithme arrive avec succès à capturer la relation non-linéaire entre les évaluations des auditeurs, basées sur une note de cinq points et certaines caractéristiques du signal transmis sur le réseau prenant en compte les paramètres mentionnés ci-dessus. Le résultat est un réseau neuronal inter-subjectif qui pourrait efficacement évaluer la qualité du son transmis sur le réseau afin de prendre d'éventuelles décisions sur l'ingénierie du réseau ou sur l'applicatif.

1 Introduction

Il y a grande demande pour l'évaluation de la qualité sonore dans les applications distribuées. En dépit de l'importance de cette demande, très peu de méthodes sont disponibles. D'ailleurs, les quelques contributions dans ce domaine se concentrent principalement sur la différenciation des algorithmes de codage sans tenir compte des paramètres de réseau [2]. Plusieurs paramètres affectent la qualité du son transmis sur un réseau, tels que la cadence de perte, le retard de bout en bout, la variation du retard (gigue), la taille des paquets, la distribution des pertes...[2][3][11][16]. Dans les sessions interactives bi-directionnelles, l'écho, les effets d'interférence et le nombre de sources de participants sont également parmi ces paramètres [16].

La qualité sonore n'est pas linéairement proportionnelle à la variation des paramètres du réseau. D'ailleurs, la variation de ces paramètres n'est même pas prévisible, et est fortement non linéaire [3][11]. La détermination de la qualité est, donc, un problème complexe, et il n'a pas été possible de le résoudre en développant des modèles mathématiques qui incluent les effets de tous ces paramètres.

En général, l'évaluation sonore est effectuée par des méthodes objectives ou subjectives. D'une part, les méthodes objectives mesurent la qualité basée sur les analyses mathématiques qui comparent les échantillons initiaux et décodés. Le rapport signal-bruit (SNR), la déformation d'Itukura-Saito, le taux de probabilité logarithmique, le rapport qualité bruit segmentaire et les mesures perceptuelles de qualité de la parole (PSQM) sont parmi les méthodes objectives [5][6][12]. Cependant, pour vérifier l'exactitude de ces méthodes, il est habituellement nécessaire de les corrélérer avec des résultats obtenus par des essais subjectifs de la qualité sonore.

D'autre part, dans l'évaluation subjective de qualité, les méthodes mesurent l'intelligibilité de la parole (il faut toujours distinguer entre la qualité du son en général et celle de la voix en particulier qui n'ont pas les mêmes critères), ou la qualité globale perçue. Les essais d'intelligibilité incluent le test modifié de rime (MRT) et le test diagnostique de rime (DRT); tandis que, dans les méthodes globales de qualité, le plus généralement utilisé pour l'évaluation de la qualité sonore est le score moyen d'opinions (MOS), recommandé par le CCITT [4]. Il consiste à permettre à n sujets d'écouter les sons afin d'évaluer leur qualité, selon une échelle de 5 points (excellent, bon, passable, mauvais, et très mauvais). Dans [17] les auteurs étudient les deux approches pour évaluer la qualité de systèmes sonores et visuels sujet à la variation des paramètres du réseau (QoS), et constatent que tous les deux donnent des résultats corrélés mais difficilement exploitables. Dans [10], les auteurs présentent une méthode de contrôle temps réel de la qualité sonore (NMR-DART) comportant des principes psychoacoustique et PSQM. C'est une technique perceptuelle de mesure de la qualité de la parole.

Bien que les études de MOS aient servi de base à analyser plusieurs aspects du traitement des signaux, ils présentent plusieurs limitations: a) des environnements d'évaluation très rigoureux sont exigés; b) le processus ne peut pas être automatisé; c) la classification ne s'adapte pas aux contextes de classification, et aux environnements dynamiques; et d) elles sont très coûteuses en temps pour une éventuelle répétition. D'autre part, l'exécution des algorithmes objectifs est habituellement comparée aux résultats obtenus pendant les essais subjectifs, c.-à-d., à la capacité de l'humain d'évaluer la qualité sonore.

Au lieu de rechercher des algorithmes pour mesurer objectivement la qualité sonore, pourquoi ne pas établir un système hybride qui prend en compte des mesures subjectives, et dont le comportement est semblable à celui des humains qui évaluent la qualité sonore?. Dans cet article, nous adressons cette question en décrivant une méthode développant ce nouveau concept. Nous illustrons notre approche en établissant un système qui tire profit des avantages offerts par les réseaux neuronaux artificiels (ANN) [13][18] pour capturer une relation non-linéaire entre plusieurs mesures non subjectives (c.-à-d., taux de pertes de paquets, taille de paquets, nombre de paquets consécutifs perdus, et codec) de signaux sonores transmis sur un réseau et une échelle de qualité de 5 points, qui émule l'évaluation effectuée par un groupe d'humains pendant une expérience conventionnelle de MOS. Nous appelons cette relation un modèle inter-subjectif de réseau neuronal (INN).

Il est important de mentionner que depuis les années 60, les réseaux neuronaux (ANN) ont été employés pour résoudre des problèmes dans les réseaux de transmission, la plupart du temps impliquant des décisions ambiguës, et dans un environnement dynamique. Il y a eu des applications réussies, adressant toujours des problèmes difficiles à aborder avec les méthodes traditionnelles [18], qui vont de l'utilisation des ANN en tant que supprimeurs d'écho adaptatifs et égaliseurs de ligne adaptatifs sur les canaux de téléphone [19], au développement de la contrôle complexe de trafic dans des réseaux ATM [15]. Néanmoins, ces problèmes tombent dans un des catégories suivantes: a) classification de pattern, b) prévision; et c) commande et optimisation. Dans notre cas particulier, l'évaluation de la qualité sonore basée sur des données produites par une expérience de MOS peut être abordée comme un problème de classification de pattern.

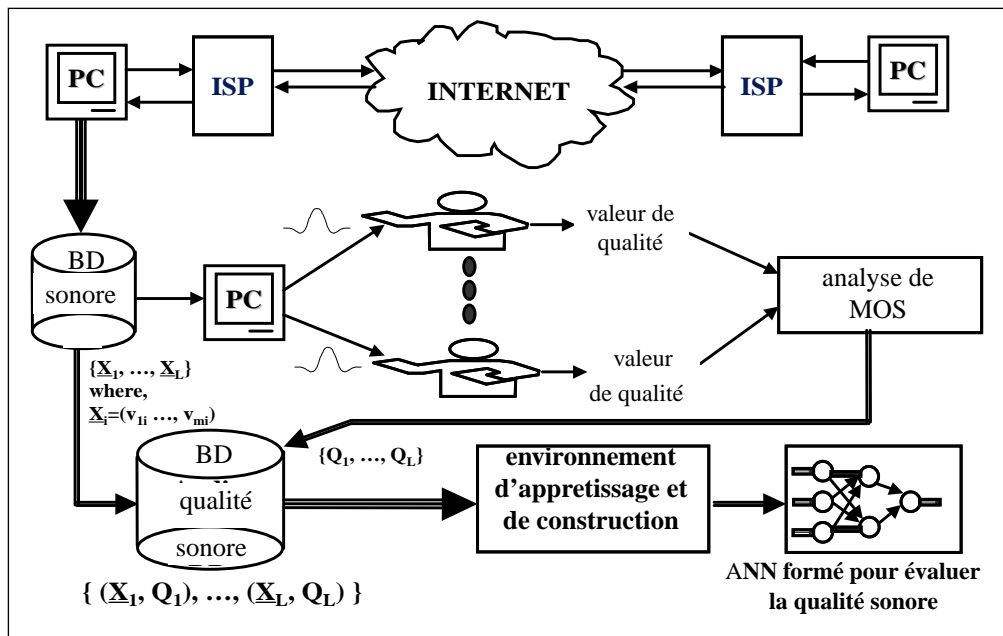


Figure 1 : Architecture globale et suite d'apprentissage

Cet article est organisé comme suit: dans la section 2, nous donnons une vue d'ensemble de notre approche pour développer un système hybride pour évaluer pertinemment la qualité des signaux sonores transmis par un réseau de paquets; la section 3 contient une brève introduction aux réseaux neuronaux artificiels; dans la section 4, nous présentons le procédé global employé pour produire des différentes bases de données sonores pour l'expérience de MOS, basées sur un banc d'essai spécifique de réseau; dans la section 5, le test et les résultats de MOS sont décrits; dans la section 6, le développement du modèle INN est décrit, aussi bien que l'analyse de son exécution; et, en conclusion, dans la section 7, quelques conclusions et futures directions de notre travail sont tracées.

2 Description de la méthode

Notre approche pour évaluer la qualité sonore en utilisant les ANNs est décrite sur la figure 1. Les paquets partent sur un réseau jusqu'à ce qu'ils arrivent à un terminal, où les valeurs d'un ensemble de variables (X_i) sont déterminées. Ces variables sont associées aux changements des paramètres réseau qui affectent la qualité. Nous supposons qu'en communications à sens unique, la qualité est sûrement affectée par la cadence de perte de paquets, au retard de bout en bout résultant des processus de stockage, à la gigue, à la taille de paquets, à la distribution et forme de pertes, et aux algorithmes de codage employés pour comprimer la transmission. Ces paramètres peuvent être divisés en familles, et différents paramètres pour des algorithmes d'encodage pourraient être identifiés, comme la fréquence d'échantillonnage, le nombre de bits par échantillon, le débit binaire d'encodage (linéaire, prédictif, etc.). Dans les sessions interactives (voix dans les deux sens), l'écho résultant des longs temps d'aller-retour (RTTs), des effets d'interférence quand l'expéditeur et le récepteur parlent simultanément, et le nombre de sources participant en cas de sessions de multicast, peuvent être considérés parmi ces paramètres affectant la qualité sonore.

Dans le récepteur de l'outil de mesures, il y a un banc d'essai adapté à l'évaluation des paramètres identifiés. Le matériel doit être choisi, et différents intervalles et valeurs des paramètres devraient être définis, aussi bien que les combinaisons de paramètres qui couvrent un éventail de situations possibles dans le réseau. Puis, en actionnant ce banc de tests, nous produisons et enregistrons un ensemble de signaux qui ont subi les déformations (BD sonore), avec leurs paramètres correspondants (X_L).

Une fois que les échantillons sont rassemblés, un essai subjectif approprié d'écoute devrait être effectué. Il doit suivre la recommandation de l'UIT-T P800 pour la transmission de la parole, et BS.562-3 [7] pour l'acoustique en général. Pour les sessions multicast il est nécessaire de suivre l'ITU-R BS 1116-1. L'essai subjectif est effectué avec n sujets, et le MOS (Q_T) est calculé, et enregistré en association avec les paramètres. Les sujets devraient écouter les signaux de telle manière qu'ils ne puissent établir aucune relation entre les échantillons et les valeurs de paramètres. La base de données résultante (BD qualité) est employée pour établir l'application d'ANN (les détails de ce procédé sont décrits dans la section suivante). Le procédé global devrait être répété, selon les besoins, pour améliorer l'exactitude d'ANN en évaluant la qualité sonore. Plusieurs ANN ont pu être combinés pour traiter différents scénarios de transmission. Une fois qu'une configuration stable de réseau neuronal est obtenue, l'architecture et les poids d'ANN peuvent être extraits afin de construire un outil simple comportant deux parties: une rassemblant les valeurs de paramètres d'entrée, et l'autre pour déterminer la qualité sonore correspondante. Bien que les ANN aient la propriété de généralisation, ils auront une performance relative à l'information fournie dans les phases d'apprentissage.

3 Les réseaux neuronaux artificiels comme classificateurs de pattern (configuration)

Un ANN est un processeur associatif distribué parallèle, comportant des éléments multiples (modèles neuronaux) fortement interconnectés. Chaque neurone effectue deux opérations: d'abord, un produit intérieur, $\underline{w}_{ij} \cdot \underline{x}_j$ d'un vecteur d'entrées \underline{x}_j , et d'un vecteur de poids, \underline{w}_{ij} , où le vecteur de poids représente les

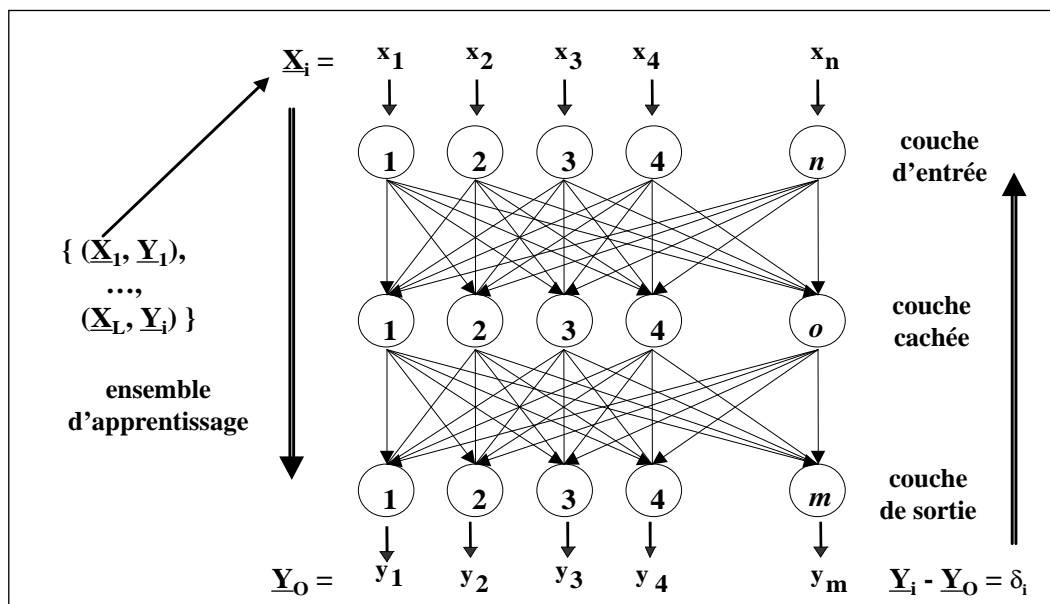


Figure 2 : Modèle de réseau neuronal utilisé

efficacités associées aux connexions venant d'autres neurones et/ou des entrées externes; et en second lieu, une relation non linéaire entre le produit interne et une grandeur scalaire, $k_i = f(w_{ij} \cdot x_j)$, où, normalement, f est une fonction non linéaire, non décroissante, continue (par exemple tanh). En construisant un ANN une architecture et un algorithme d'étude doivent être choisis. Il y a des architectures multiples (par exemple, réseaux multicouche dits "feedforward", réseaux récurrents, réseaux bi-directionnels, etc.), aussi bien que des algorithmes d'étude (algorithme par exemple, de retro-propagation, de Kohonen LVQ, algorithme de Hopfield, etc.) [18].

Comme classificateurs de patterns, le rôle d'un ANN est de rechercher une fonction non linéaire qui lie un ensemble de vecteurs d'entrées (des patterns ou configurations), X_n , avec leurs vecteurs correspondants de sortie (catégories), Y_n . Cette liaison est établie en extrayant l'expérience incluse dans un ensemble d'exemples (ensemble d'apprentissage), en suivant un algorithme d'apprentissage. Ainsi, en développant une application avec les ANN, il faut d'abord un ensemble d'exemples connus par N représenté en termes de configurations et catégories (c.-à-d., dans paires (X_n, Y_n) , où $n=1,2, \dots N$); en second lieu, une architecture appropriée devrait être définie; et, en troisième lieu, un algorithme d'apprentissage doit être appliqué afin d'établir la relation.

Des relations fortement non linéaires peuvent être obtenues en utilisant l'algorithme de retro-propagation pour l'apprentissage et l'adaptation, et un réseau neuronal de "feedforward" de trois-couche se composant d'une couche d'entrée, d'une couche cachée, et d'une couche de sortie [13]. Dans cette architecture (voir la figure 2) : les entrées externes sont les entrées pour les neurones dans la couche d'entrée; les sorties scalaires de ces éléments neuronaux dans la couche d'entrée sont les entrées des neurones dans la couche cachée; et, les sorties scalaires dans la couche cachée deviennent, les entrées pour les neurones dans la couche de sortie. En appliquant l'algorithme de rétro-propagation, tous les poids des valeurs initiales sont choisis aléatoirement. Puis, pour chaque paire, (X_n, Y_n) , dans la base de données, le vecteur X_n (c.-à-d., le pattern) est placé comme entrée de la couche d'entrée, et le processus est effectué par la couche de sortie en passant par la couche cachée. Une réponse est produite. Une erreur d est calculée en comparant le vecteur Y_n (classification) à la réponse de la couche de sortie, Y_{on} , à la configuration X_n . S'ils diffèrent (c.-à-d., une configuration est mal classifiée), les valeurs de poids sont modifiées dans tout le réseau en conséquence selon la règle généralisée de delta:

$$\underline{w}_{ij}(t+1) = \underline{w}_{ij}(t) + \eta \delta x_j$$

où, \underline{w}_{ij} est le poids pour la connexion que le neurone i , dans une couche donnée, reçoit du neurone j de la couche précédente; x_j est la sortie du neurone j dans cette couche; δ est un paramètre représentant la cadence d'apprentissage; et d est une mesure d'erreur. Dans le cas de la couche de sortie, $d = Y_n - Y_{on}$; considérant que dans toutes les couches cachées, d est une erreur estimée, basée sur la rétro-propagation des erreurs calculées pour la couche de sortie (pour plus de détails [13][18]). De cette façon, l'algorithme de rétro-propagation réduit au minimum l'erreur globale associée à toutes les paires, (X_n, Y_n) , $n=1,2,\dots N$, dans la base de données. Le processus d'entraînement continue jusqu'à ce que toutes les configurations soient correctement classifiées, ou une erreur minimum prédéfinie a été atteinte. En conclusion, si des exemples contenant de nouvelles informations arrivent, les ANN peuvent être recyclés pour s'adapter à la nouvelle situation; c'est-à-dire, les algorithmes d'apprentissage donnent aux réseaux neuronaux l'avantage d'une adaptabilité élevée, qui leur permet d'optimiser leur exécution dans des environnements dynamiques.

4 Description de l'environnement de test et collecte des paramètres.

Pour valider notre approche nous avons considéré les expériences de MOS dans le cadre de transmission de la parole sur un réseau IP sans qualité de service. Nous avons choisi un certain nombre de paramètres réseau qui ont un effet dominant sur la qualité de la parole. Ces paramètres sont: a) perte moyenne de paquets; b) taille de paquets; c) nombre de paquets consécutifs perdus; et d) l'algorithme de codage. En outre, le retard et la gigue sont également pris en considération mais inclus dans le paramètre général de taux de pertes. Ceci est confirmé par la plupart des outils "audio" de l'internet [16]. Nous avons considéré seulement le cas de la session à sens unique. Nous avons utilisé une séquence audio normalisée par l'UIT employés pour tester l'algorithme G726 [9] de codage. Elle contient 3000 échantillons de phrases en fran-

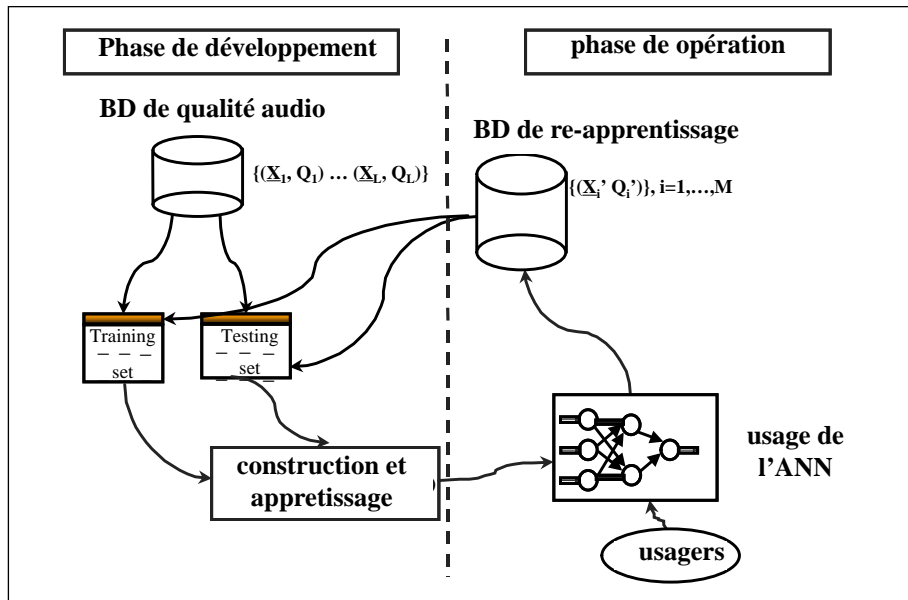


Figure 3 : Environnement de simulation pour construire l'ANN

çais, dans le format PCM. Les cinq premières phrases ont été utilisées avec une moyenne de 4 mots par phrase. Les valeurs des paramètres ont été rassemblées sur un réseau IP se composant d'un émetteur, d'un routeur et d'un récepteur.

L'émetteur contrôle la taille de paquets et l'algorithme de codage. Il emploie le protocole de RTP pour contrôler la transmission. Le routeur contrôle le taux de pertes et le nombre de paquets consécutifs perdus pendant une session donnée. Le récepteur décode et enregistre les paquets reçus de la parole, les paquets détruits sont remplacés par des périodes de silence (ils pourraient également être remplacés par un bruit de confort) et détermine les pertes (grâce entre autres au RTCP).

Nous avons utilisé des tailles de paquet de 20, 40 ou 80 millisecondes de parole, par le codage PCM linéaire, ou GSM comme algorithmes. Nous avons configuré le routeur avec un taux de pertes allant de 0 à 50%, tandis que le nombre de paquets consécutifs perdus dans des rafales est de 1 à 5 paquets. À la destination, les pertes du protocole de RTP sont déduites (en utilisant des estampilles et/ou des numéros de séquence). Nous avons choisi 65 scénarios qui incluent les paramètres précédents. Nous avons ensuite fournis les résultats sonores dans un ordre aléatoire à l'audience afin d'effectuer le test MOS.

La phase d'évaluation est conforme à la recommandation de MOS UIT-T P-800 [4]. Un groupe de sept sujets ont été invités à écouter 5 phrases, chacune avec 4 mots en moyenne, et à évaluer leur qualité en s'accordant avec une échelle MOS de 5-point. La correspondance entre la note de la parole et les paramètres ont été complètement dissociées afin d'avoir une évaluation très neutre. Les résultats obtenus sont en accord avec d'autres expériences dans la littérature [17], par exemple, la qualité sonore se dégrade à mesure que la cadence de perte de paquet augmente.

5 Un réseau neuronal inter subjectif

Les phases de la conception de l'ANN pour des tâches de classification de patterns, sont montrées sur la figure 3. Juste comme dans la programmation, il y a une phase de développement et une phase de mise en oeuvre. Pendant la phase de développement, une base de données devrait être recueillie, contenant des exemples connus de configurations, $X_j, j=1 \dots L$, et leur classifications correspondantes, Q_j ; dans notre cas particulier, ceci a été fait en entreprenant une expérience de MOS (voir la section 4). Puis, cette base de

données est divisée en deux jeux: un pour la formation, et l'autre pour le test. Ensuite, l'architecture de l'ANN doit être définie, aussi bien que l'algorithme approprié d'étude pour l'entraîner. La phase de développement est finie quand l'ANN opère à un niveau de précision pré-établi, ou toutes les configurations de l'apprentissage et de test sont classifiées correctement, ou une erreur globale minimale a été atteinte. Une fois que l'ANN est en opérationnel, son exécution devrait être surveillée. Des cas d'anomalie, comme les cas de grand succès "hit", devraient être enregistrés comme paires de configurations, X_i , $i=1...M$, et des classifications, Q_j à chaque fois que l'exécution de l'ANN va en-dessous d'un seuil prédéfini doivent être enregistrées et, l'ANN devrait être recyclé. En ajoutant les exemples recueillis pendant la phase d'exécution, il est nécessaire de tester les résultats obtenus. Quand toutes les configurations marchent correctement ou une erreur globale minimale est atteinte, le nouvel ANN est prêt. Ce processus peut être automatisé, donnant à l'ANN la caractéristique de l'adaptabilité; c'est-à-dire, les ANN peuvent améliorer leur exécution pendant en s'entraînant avec de nouvelles données.

6 Construction du réseau neuronal

Afin d'établir un dispositif d'évaluation de la qualité sonore, Q_j , d'un ensemble de signaux, X_j , $j=1...L$, ayant subi des pertes sur un réseau, un ANN sera utilisé pour créer un modèle qui associe des variables avec une qualité sonore de 5 points. Une architecture neuronale "feedforward" de trois-couche, et l'algorithme d'apprentissage de rétro-propagation ont été choisis. Le nombre de paramètres réseau qui affectent la qualité sonore définissent le nombre de neurones dans la couche d'entrée (4 dans notre cas particulier), alors que la couche de sortie était complétée par un neurone représentant le MOS, dont les valeurs s'étendent entre 1 et 5. Un simulateur commercial (version 2,0 de prédicteur de Neuroshel pour MS-Windows) a été utilisé pour conduire l'apprentissage. Plusieurs séries de nombres de neurones dans la couche cachée ont été testés. Pendant les simulations les ANN qui donnaient les meilleurs résultats étaient entre 44 et 59 neurones dans la couche cachée. Dans toutes les couches des fonctions "sigmoïdes" ont été utilisées.

La base de données de "MOS" a été divisée en deux parties: un groupe représentant 50 échantillons des résultats; et l'autre groupe représentant les autres 15 échantillons des résultats. En utilisant le premier groupe (la base de données d'apprentissage), un ANN a été construit, et les résultats obtenus sont montrés dans la figure 4. Nous avons obtenu: un coefficient de corrélation de 0.99514, et une erreur de 0.057 en moyenne; c'est-à-dire que le modèle neuronal inter-sujetif s'adapte tout à fait à la manière avec laquelle les humains ont évalué la qualité de la parole.

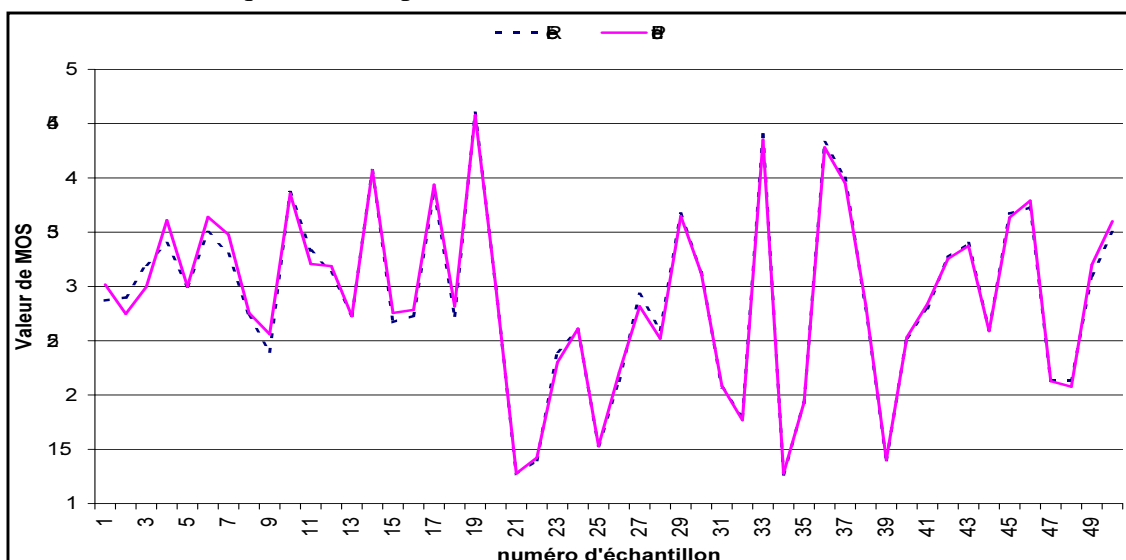


Figure 4 : Les valeurs réelles contre celles prédites de MOS pour la base de données d'apprentissage.

Comportement de l'ANN.

Afin de répondre à la question : à quel point peut on se fier à un réseau à neurones, nous avons appliqué l'autre groupe de résultats (la base de données de test). Les résultats étaient : un coefficient de corrélation de 0.98953, et une erreur de 0.09 en moyenne. De nouveau les performances de l'ANN étaient excellentes, comme on peut l'observer dans la figure 5.

D'ailleurs, de ces résultats, nous pouvons voir que notre méthode peut être utilisée dans des réseaux avec un comportement dynamique. Une fois que l'ANN est bien formé, il peut interpoler les résultats même si les entrées ne sont pas nécessairement des échantillons employés pour le former. En d'autres termes, l'ANN peut imiter la voie avec laquelle les sujets humains peuvent estimer la qualité sonore subjective.

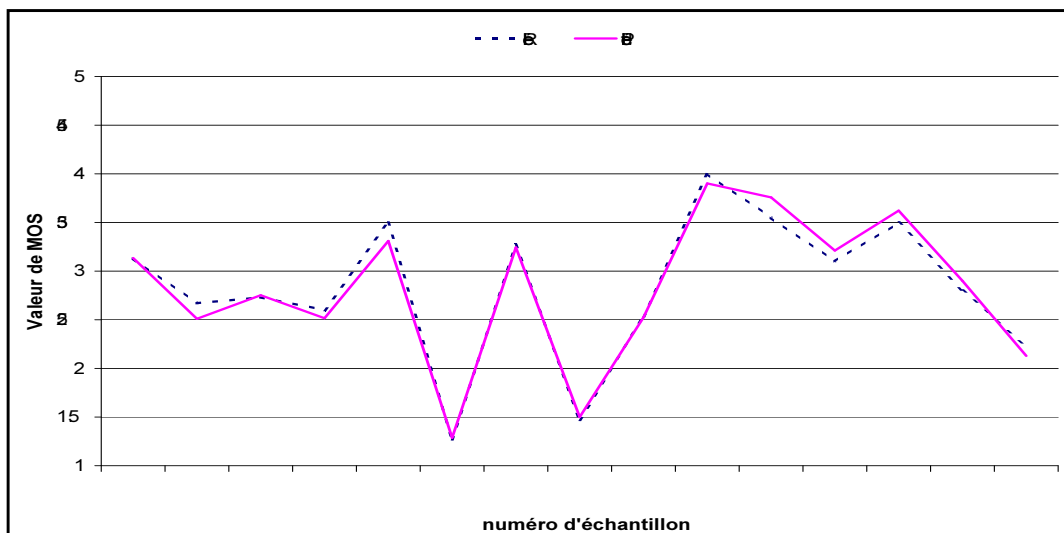


Figure 5 : Les valeurs réelles contre celles prédites de MOS pour la base de données de test.

7 Conclusion

Nous avons présenté une méthode basée sur les réseaux neuronaux permettant d'évaluer la qualité sonore de communications transmises sur l'Internet en temps réel. Cette méthode prend en compte un nombre de paramètres que nous estimons être très pertinents dans le résultat définitif de la qualité de son. Nous avons effectué une procédure d'évaluation des échantillons sonores selon les normes de l'UIT-T et ensuite procédé à l'apprentissage du réseau à neurones. Ensuite, nous avons vérifié la bonne réaction du nouveau réseau neuronal à de nouveaux échantillons. Les résultats sont extrêmement encourageant et il est maintenant possible de combiner ce nouvel outil à une procédure de décision pour l'adaptation du flux aux contraintes du réseau.

8 Références

- [1] Kostas, T., Borella, M., Sidhu, I., Schuster, G., Gabrieç, J., and Mahler, J. "Real-time voice over packet-switched networks". IEEE Network, January/February (1998).
- [2] Bolot, J.C, Fosse-Parisis, S., and Towsley, D. "Adaptive FEC-based error control for Internet telephony". Proc. Infocom'99, March (1999).
- [3] Bolot, J.C. "Characterizing end-to-end packet delay and loss in the Internet". J. of High Speed Networks, Vol. 2, No. 3 (1993).
- [4] CCITT, Recommendations of the P Series. "Method for the evaluation of service from the stand point of speech transmission quality". CCITT Red book, Vol. V-VIIIth Plenary Assembly (1984).
- [5] J. Deller, J. Proakis, J.H.L. Hansen, "Discrete-Time Processing of Speech Signals", MacMillan Series for Prentice-Hall, New York, NY, 1993.
- [6] J.H.L. Hansen, B.L Pellom, "Speech Enhancement and Quality Assessment: A Survey", submitted to IEEE Sig. Proc. Mag. Nov. 1998.
- [7] ITU-R, Rec. BS.562-3. Subjective assessment of sound quality.
- [8] ITU-R, BS.1116-1 Methods for the subjective assessment of small impairments in audio systems including multichannel sound systems
- [9] ITU-T Recommendation G.726 - 40, 32, 24, 16 kbit/s Adaptive Differential Pulse Code Modulation (ADPCM), (1990).
- [10] OPTICOM OPERA™ -"The new generation of measurement systems to analyze the perceived audio quality of speech and music codecs. <http://www.opticom.de/>
- [11] Paxons, V. End-to-End Internet Packet Dynamics. IEEE/ACM Transactions on Networking, 7(3) pp. 277-292, 1998.
- [12] S.R. Quackenbush, T.P. Barnwell, M.A. Clements, "Objective Measures of Speech Quality", Prentice-Hall, NJ, 1988.
- [13] Rumelhart, D.E., Hinton, G.E., and Williams, R.J. "Learning internal representations by error propagation". In: Parallel Distributed Processing, Vol. 1. The MIT Press, Cambridge, Mass. (1986).
- [14] S. Rety. Développement d'une passerelle téléphonique RNIS/H.323. Technical Report.
- [15] Special Issue on: Neurocomputing in High Speed Networks. IEEE Communications Magazine Vol.33, No.10 (1995).
- [16] RAT: Robust Audio Tool. <http://www-mice.cs.ucl.ac.uk/multimedia/software/rat/publications.html>
- [17] Watson, A., and Sasse, M.A. "Measuring Perceived Quality of Speech and Video in Multimedia Conferencing Applications. Proceedings of ACM Multimedia'98, 12-16 September, 1998. Bristol, England pp 55-60.
- [18] Widrow, B., Rumelhart, D., and Lehr, M. "Neural Networks: Applications in Industry, Business and Science". Communications of the ACM, Vol. 37, No. 3 (1994).
- [19] Widrow, B., and Stearns, S. "Adaptive signal processing". Prentice-Hall, Englewood Cliffs, N.J., (1989).