

Chapter 7. Field Test Results

Section 7.1. NBC Demonstration

Once the coding method had been established, and laboratory testing had validated the concept, the project sponsors hoped to raise additional capital to support further development. They began scheduling demonstrations in an attempt to lure other investors into infusing supplementary funds. Many of these demonstrations occurred in the DSP Research Lab (DSPRL) or the CWT Lab. However, while trying to entice NBC, the nation-wide television network, the sponsors desired a tape containing several NBC commercials modified with code insertions. After such a tape was created, and after computers in the DSPRL confirmed successful detection of the concealed codes, the tape was sent to NBC. The sponsors wanted to convince NBC's executives and engineers that such hidden audio coding was feasible. Since this particular demonstration pre-dated by about a year even the earliest AudioLink prototype, the DSPRL computers and soundcards had to be used to validate code reception.

When asked to produce this tape, the first order of business was acquiring the commercials. From the beginning of the project we had been using audio signals sampled from television commercials by a computer soundcard, but here the corresponding video was also requested. This was beyond the capabilities of our laboratory hardware, and so Virginia Tech Video Broadcasting Services (VBS) was asked to assist. They acquired nine NBC commercials from satellite transmissions, and stored them to a high quality tape. From there they sampled the audio of each

commercial with a computer and soundcard, and posted the resulting WAV files on a File Transfer Protocol (FTP) server connected to the internet. We downloaded the WAV files to our local area network in the DSPRL, and began the code insertion process.

Since the thirty-second CD-quality audio files were far too large to load into Matlab in their entirety, we first split them up into manageable segments of about five seconds each. CD-quality implies two channels, a 44.1 kHz sampling rate, and a sample size of sixteen bits. Here only monaural signals were used, although our codes could also be inserted into stereo signals. Each thirty second commercial thus contained about $44,100 \text{ samples per second} * 2 \text{ bytes per sample} * 30 \text{ seconds} = 2,646,000 \text{ bytes}$. Since Matlab increases the sample sizes to eight bytes each, more than ten megabytes would be required just to hold the original copy of the signal, never mind intermediate and final versions. Therefore operating on smaller five-second segments in sequence was a necessity.

Using the steps and procedures outlined in Section 3.2.3, we searched for candidate code insertion locations. Each candidate position was tested, and the best targets were chosen by trial and error. Once the locations were set, the relative power levels were also chosen empirically. The modified audio segments were then saved back to WAV files, and were recombined with the other segments to re-form the full thirty-second commercials. The modified commercials were down to the number of samples the same length as the originals. These modified commercials were transmitted by FTP back to the VBS server. VBS personnel then combined the modified audio with the saved video by synchronizing and mixing the two, and saved the results to another high

quality tape. Copies of this tape were provided to us for experiments and demonstrations, and a copy was also sent to NBC's headquarters in New York.

The above description neglected one crucial step in the commercial modification process. Although quartz crystals and divide-by factors control the sampling rates of computer soundcards, the crystals and factors can vary from one card to the next. Therefore, two soundcards can sample at drastically different rates even when both are nominally set to the same sampling frequency. Sound card crystals are typically stable to about 300 parts per million, or 0.03 percent, and so frequency variation is not the problem. Rather the difficulty is that there can be a bias in the rate itself. For instance, although both the VBS computer and the DSPRL code insertion computer ostensibly sampled and played back at 44.1 kHz, their actual rates differed by almost six percent. Since the IVDS communication scheme relies on sinusoids at precise frequencies (or slightly shifted versions of them), a six-percent deviation completely prevents successful transmission and reception of the codes.

In order to compensate for differing sampling rates, we created and incorporated what we dubbed a Sampling Rate Conversion Factor (SRCF), as discussed in Section 5.6.2. This SRCF changed the normalized frequencies of the sinusoids, so that their actual frequencies would be correct when played back on a particular computer through a particular soundcard. Therefore, in the commercial modification process, a trial run had to be performed first so the necessary SRCF could be computed. This was achieved by examining where the frequency locking sinusoids landed in the frequency domain, relative to where they should have been. The SRCF is then simply the inverse ratio of the

centers of the two frequency locations. Subsequent code insertions would then use this factor whenever that same computer was responsible for the audio sampling and playback.

The NBC commercial demonstrations were extremely successful. In two of the nine commercials even the trained audio engineers at VBS could not find the code insertion points after repeated listening episodes. And in the other seven commercials the insertions were imperceptible to most observers, although a few of the young observers with sensitive ears could identify them. In no cases were the codes described as objectionable. If the reader is viewing an electronic version of this document, two examples of the modified commercials can be heard by clicking on the icons below:



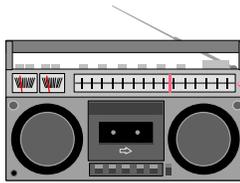
Reports from the NBC meetings suggested that the executives were impressed at how imperceptible the modifications to the audio were, despite being detectable by our computer hardware. In the end, however, NBC declined to participate in our project. Since their engineers had advised them in the past that such coding was impossible, they feared investing in a pursuit that might not be successful in the long term. They would be vulnerable to criticism if they ignored past advice and spent money on a project that eventually failed.

Section 7.2. Virginia Tech Orange Bowl National Broadcast

After their successful 1996 season, Virginia Tech's football team was invited to compete against the University of Nebraska in the nationally televised Orange Bowl. As is customary, Virginia Tech was given a thirty-second commercial spot to advertise the university. The IVDS sponsors convinced Virginia Tech's administration to allow a code insertion into the commercial, so our system could be tested during a national television broadcast. The Orange Bowl provided such a unique testing opportunity that all of our project resources were diverted to meeting the deadlines for the university's production.

The code insertion procedures for the Virginia Tech promotional commercial were the same as used for the NBC commercials. However, we found that finding a suitable target location for the code was extremely difficult for this particular commercial. After investigating the situation, we discovered the reason for the complication. The background music in the commercial contained steady tones and chords, which changed relatively infrequently, and which were constant between changes. During the code insertion process, bandstop filtering removes all of the energy in the original audio that occurs at the same time and at the same frequencies as the code sinusoids (refer back to Section 3.2.2 for details). Inevitably, regardless of temporal placement, this filtering eliminated some of the musical tones, creating a discontinuity that was very perceptible. It was not the code energy that became noticeable, but rather it was the disruption in the background music that was disconcerting. In the end we had to settle on placing the code within the first second of the commercial, as the music begins

and rises toward its steady state quality. The code was imperceptible to most observers in this location, but there was a danger that the commercial could be slightly clipped during broadcast. If such clipping occurred the code segment could be lost. If the reader is viewing an electronic version of this document, the Virginia Tech Orange Bowl commercial can be heard by clicking on the icon below:



During the Orange Bowl on the evening of December 31, 1996, several prototype AudioLinks were placed on televisions at test stations in anticipation of the commercial. In some cases the television signal was received by a local broadcast from a local station. In others it was received by cable, where the cable companies initially obtained the signals from a satellite link. At another site the signal was acquired by a direct digital satellite broadcast. The AudioLinks successfully detected the codes despite the various transmission paths. Needless to say we were all elated by the results, especially after the many hours of hard work expended in setting up the experiment. This success also provided the sponsors with results they could cite when courting other potential investors.

Section 7.3. Hotel Roanoke Field Test

On October 22nd and 23rd, 1997, the CWT hosted a “Wireless Opportunities Workshop (WOW)” at the Hotel Roanoke in Roanoke, Virginia. The IVDS sponsors decided that the conference provided an outstanding opportunity to conduct a complete

field test of our prototype system. This field test included AudioLinks in user hotel rooms (and a few in the hotel's restaurant/bar), a repeater unit located in a central room, and a host computer connected by a serial cable to the repeater unit. With the hotel's cooperation, we were allowed three channels on their local cable distribution system for use in the experiment. We produced our own programming to place on the three channels. It was decided that one channel would be used by the CWT to promote the WOW conference and workshop, and another would contain commercials for the hotel and for various local merchants and restaurants. The third channel would be used exclusively for an interactive game of trivial pursuit.

For the WOW channel the CWT staff produced 58 PowerPoint slides advertising the various companies, speakers, and sessions of the conference. Background music was selected and was sampled from a CD and saved to WAV files. We created a master presentation in PowerPoint that would call up the 58 slides of the sub-presentations in the correct order and with correct timing and transitions between slides. The WAV files were ordered and assigned to the numerous slides, with each WAV file persisting over several slides' duration, and codes were inserted wherever necessary. The PowerPoint presentations were then played back in automatic slideshow mode, and the video output was sent through a scan converter to create an NTSC signal. This NTSC signal, along with the audio output from the computer soundcard, were input to a VCR which then recorded the commercials onto VHS tapes.

Each code in the audio identified the corresponding slide and specified what action the AudioLink was to take upon reception of the code. All of the codes inserted

for the WOW conference channel contained a digital command of two. As described in Section 5.5 this command causes the AudioLink to light its LED at the front of its case to indicate that a code has been received. The device then waits for an infrared response from the remote control. If a valid response is received, the AudioLink combines its serial number (identifying the user), the digital message from the commercial (identifying the commercial), the infrared key press (identifying the user's response), and a time-stamp (identifying the time of key press), to form a message to be transmitted to the repeater unit. Once the host computer receives the composite message from the repeater, appropriate action is taken. In this case a packet of information about whatever entity the particular slide was advertising (say, a conference speaker) would be sent to the user's hotel room. An example slide from this channel is provided below in Figure 7.1.

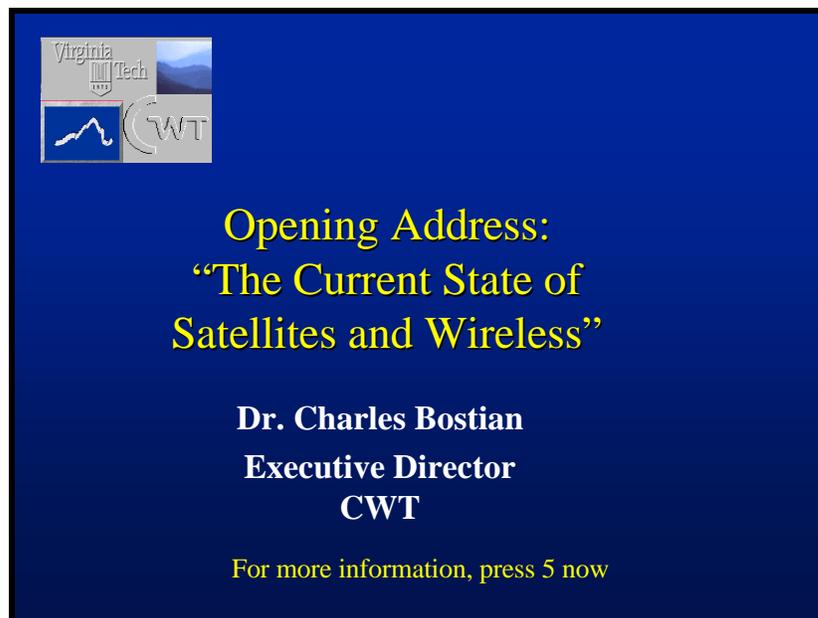


Figure 7.1 Example Slide from the WOW Conference Channel

The WOW channel's slides contained visual prompts for user interaction, and because of the use of digital command two the AudioLink did not verbally prompt the user. This mode of operation demonstrates what is possible once AudioLinks are prevalent among television viewers. Such visual prompting also provides good advertising for the IVDS system, since viewers without AudioLinks will want to know what they are and what they can do, and will possibly request or purchase one for themselves. It should be noted that no effort was made to hide the codes for the WOW channel. The codes were simply inserted at convenient times. The rationale for this modus operandi was that the WOW channel's purpose was to advertise the IVDS system, and so code concealment was not relevant.

The second channel, the "commercial" one, also contained PowerPoint slides. The Marketing Club of Virginia Tech was hired to help the CWT staff produce slides for the nineteen merchants and restaurants that participated in the experiment. The companies who participated received free advertising in exchange for their cooperation. Most of them supplied coupons, directions, menus, etc. to be given to users who requested additional information when prompted by the AudioLink. Each commercial consisted of between two and six slides, and all of the sub-presentations were again called from a master presentation in PowerPoint. The master presentation controlled the timing and transitions between slides, and played the appropriate WAV files at the appropriate times, as was done for the WOW channel. An example slide from the commercial channel is shown below in Figure 7.2.



Figure 7.2 Example Slide from the Commercial Channel

For the commercial channel, some of the advertisements contained music, some contained narration, and some contained a mixture of the two. An undergraduate student named Mannin Dodd, who was hired to help with the IVDS project, did much of the speech narration. His voice was sampled in the acoustically damped room on the fifth floor of Whittemore Hall and saved to WAV files. After filtering and cleaning, his voice was usually then mixed with selected background music. The audio signals often required fading in and out, and other similar editing effects. A code was then inserted into the final result. All audio signal processing was performed in the DSPRL.

Each inserted code that was broadcast over the commercial channel specified a digital command of four. As described in Section 5.5 this command causes the AudioLink to light its LED at the front of its case to indicate that a code has been received, and a verbal prompt of “Press a key for more information” is output through its

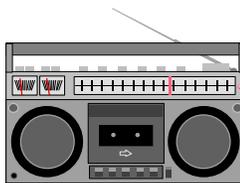
speaker. The device then waits for an infrared response from the remote control. If a valid response is received, the AudioLink combines its serial number (identifying the user), the digital message from the commercial (identifying the commercial), the infrared key press (identifying the user's response), and a time-stamp (identifying the time of key press), to form a message to be transmitted to the repeater unit. Once the host computer receives the composite message from the repeater, appropriate action is taken. In this case a packet of information and other materials (say, coupons) for whatever entity the particular slide was advertising (say, a restaurant) would be sent to the user's hotel room.

Unlike the WOW channel, the commercial channel's slides did not contain visual prompts for user interaction. But because of the use of digital command four, the AudioLink verbally prompted the user instead. This mode of operation demonstrates what is possible in the early stages, when AudioLinks are still rare among television viewers. The television video programming is not modified at all, and the code insertions are relatively if not perfectly imperceptible. Therefore viewers without AudioLinks would not know that a code was inserted or perhaps that AudioLinks even existed. This is comforting to advertisers who do not want serious modifications made to their commercials. Later, when the devices become more popular, the video too can be used for user prompting. As mentioned above, this latter mode also helps to advertise the IVDS system.

The third channel that was broadcast during the Hotel Roanoke field test was used exclusively for an interactive game of trivial pursuit. John Deighan, a graduate student in

the Industrial and Systems Engineering Department, assisted Willard Farley in writing the game in software to be used with the IVDS system.

A game computer generated a screen containing the questions and a list of the potential answers. The video signal was intercepted by a scan converter which created an NTSC version of the screen image, and this converted video signal was inserted directly into the Hotel's cable network on the game channel. While the screen image was being displayed, the computer would also output audio signals through its sound card. These too were fed directly into the cable system. Thus the computer's audio and video output were used as the programming content for the game channel. The audio consisted of background music and voice messages, both of which were pre-recorded. The codes were used in their "raw" form (i.e., they were not hidden in another audio signal) as buzzer prompts to indicate the posting of a new question. If the reader is viewing an electronic version of this document, the raw buzzer sound can be heard by clicking on the icon below:



If a user was tuned to the game channel, the AudioLink would detect the buzzer prompt (the code) as each question was posted. Its front panel LED would indicate that it was waiting for a reply. The user's guess at the answer would be represented by the key pressed on the remote control. If the AudioLink detected a valid key press (say digits one

through four, to represent possible answers one through four on the screen), it would begin transmission of the results. The AudioLink would combine its serial number (identifying the user), the digital message from the commercial (identifying the game), the infrared key press (identifying the user's response), and a time-stamp (identifying the time of key press), to form a message to be transmitted to the repeater unit. The time stamp was used to determine the order in which the contestants answered the questions. More points were given to contestants that answered immediately (if correct), and the points scored decreased as time progressed after the question.

Once the host computer received the composite message from the repeater, appropriate action was taken. In this case a modem link from the host computer relayed the responses to the game computer. After displaying the correct answer on the screen, the game computer tabulated scores for all of the game contestants, and displayed the current rankings and scores. After a short pause, the next question was posted and the cycle repeated. The winners of each round were given a CWT tee shirt and the overall winner after the two-day trial received \$500. (The video technicians who manage the hotel's cable system played the interactive game almost incessantly, probably at the expense of their duties. They finished as the overall winners.)

The Hotel Roanoke field test was quite successful, and the sponsors were pleased, despite the appearance of a few hardware problems. Since the purpose of the field test was to identify potential problems, they did not expect trouble-free operation. The first obstacle was that the only available repeater unit failed midway through the first evening, and so the testing had to be suspended. However, it was repaired before the next

morning, and the trial was resumed. Also, the AudioLinks seemed to have difficulty detecting a few of the quieter codes on the commercial channel (where the codes were hidden). Investigations revealed that this was due to using copies of VHS tapes, and because the various commercials were not normalized in volume. Several commercials were much louder than others. As users watched their televisions, they would reduce the volume to compensate for the overbearing commercials. Reducing the television volume, however, was detrimental to code detection during the quieter commercials. Furthermore, using copies of VHS tapes instead of the originals not only introduces more noise, it also effectively widens the range of the tape wow and flutter effects.

Another reason for the detection difficulties is that there is a fundamental tradeoff inherent in the code insertion process. The better the code is hidden, the harder it is to detect. Increasing the code duration or relative volume enhances its “detectability,” but the code also becomes more noticeable (i.e., its “hideability” is reduced). In a few of the commercials we probably erred too far towards successful hiding, at the expense of detectability. During the code insertion process we did not have adequate time to properly search for the best locations, and so we had to settle for mediocre spots. In an attempt to compensate for their noticeable presence, the relative powers of the codes were reduced, probably beyond what would have been a safe level. If more time had been available, the codes could have been better hidden, with even larger relative volume levels.

Another complication that arose was false detections. Several units verbally prompted for interaction despite no codes being present. Such false detections were due

to random noise creating spectral patterns that satisfied the sinusoid criteria at a sufficient number of FFT bin locations. As mentioned in Section 3.1.3, one way to reduce the occurrences of these false detections is to require the results from the triplication code to satisfy the CRC check. Currently, if the triplication result fails the CRC test, a second chance at reception is obtained by considering the results from the three independent frequency bands. Although at times this can assist reception, it also raises the rate of false detections.

Section 7.4. Virginia Tech Campus Tests

After the success at the Hotel Roanoke, the project sponsors wanted to conduct another field test, but this time on Virginia Tech's campus. The hope was that AudioLinks could be issued to students in dorm rooms, and one or more of the campus cable channels could be used. However, the university administrators were concerned about such general distribution, since weaker students might sacrifice study time in order to participate in the experiment. Therefore only the dormitory resident advisors were invited to participate, and seven AudioLinks were distributed midway through the Spring 1998 semester. A single campus cable channel was acquired, and the same interactive trivial pursuit game was broadcast for a period of sixteen days.

The repeater unit was placed on top of Slusher tower, one of the highest points on campus. This point provided outstanding reception coverage, even when the AudioLinks were transmitting from the fourth floor of Whittemore Hall. Such transmissions are not only from the inside of a large building, but also from across campus.

This field test too was successful, and no major hardware failures occurred. Once again the AudioLinks occasionally prompted users for interaction when no codes were actually present (false detections of codes). Also, at times the infrared codes were not identified from the first key press of the remote control. This is not a major problem, since the user is notified immediately and prompted to re-press a valid key.

One other software bug that was found involved the time stamping of the remote control key presses. Once a question is posted to the TV screen, a user can press an invalid (non-numeric) key on the remote control. The AudioLink then verbally informs the user that an invalid key was pressed, and instructs the user to press a valid key. However, the time stamping occurs at the moment of the initial key press. Therefore, one user found that he could press an invalid key to lock in a quick response time. Then he could wait until the correct answer was displayed on the screen, and he could press a valid (and correct!) key. Not only would he get the answers correct 100 percent of the time this way, but his responses were time stamped at the moment of his initial key press. At first he just seemed like an extremely good player, but further investigation revealed his trick. This bug (or feature, depending on your point of view), can be easily fixed in software.

Although the experiment involved only a single cable channel and seven AudioLink users, the sponsors considered it a success. The participants in the experiment reportedly enjoyed having the devices and playing the games, and a few requested that the experiment extend for a longer duration. The test provided useful information, and once again demonstrated the potential of the IVDS system.

Section 7.5. Future Field Tests

Although future plans have not been finalized at the time of this writing, the PISA sponsors reportedly have found several television stations who are interested in using the AudioLinks and participating in future field tests. Two of the stations are located in Mexico, and one is in St. Louis, Missouri. If enough interest and capital can be raised, the IVDS system may yet become commonplace in Mexico and the United States.