

Technical Properties of Arbitron's PPM System

Dr. Barry Blesser

Director of Engineering, 25-Seven Systems, Inc.

August 18, 2009

Executive Summary

Arbitron's Portable People Meter™ (PPM™) system* is currently replacing their old paper diary system. There are a number of radio veterans questioning PPM's accuracy because listenership reports have frequently been quite different from those of the diary system.

After analyzing public information, I conclude that this system appears to be technically sophisticated and well conceived; it can be expected to be extremely reliable when tested under "typical" conditions. However, a more careful analysis suggests that there may be real-world scenarios that dramatically degrade PPM performance. Regardless of how much testing Arbitron performed before releasing the PPM system, the real world of thousands of radio stations exposes the system to an almost infinite variety of idiosyncratic properties of particular programs, speaking styles of announcers, and listening environments.

Radio stations that believe their listenership has been inaccurately reported by the PPM system may wish to determine if their context deviates from the assumptions upon which the PPM system is designed. Though I conducted a comprehensive search of the publicly available literature, I found no discussion of these assumptions, much less information on how to compensate for them.

We respect that, for legal and/or economic reasons, Arbitron may not be willing or able to share all the technical details of its PPM system with our industry. Our industry, therefore, needs to do its homework.

This article explores the technical design of the PPM system in relatively non-technical language so that station managers, programmers, and engineers can understand the possible implications of these "hidden" assumptions. Further, it frames the relevant questions that we should be addressing and provides suggestions for experiments to evaluate the PPM system in each station's context.

* PPM and Portable People Meter are trademarks of Arbitron, Inc.

25-Seven Systems, Inc.

1085 Commonwealth Avenue #360, Boston, MA 02215 USA
+1 (617) 789-4673 • e-mail: info@25-seven.com • <http://www.25-seven.com>
© Copyright 2009 25-Seven Systems, Inc. All rights reserved

Background and overview

I've been asked by several of my radio engineering colleagues to comment on the technical properties of the PPM audience measuring system. Inquirers were particularly interested in possible technical explanations of why some program formats on some stations have received dramatically lower audience ratings compared to that of the old diary system.

To respond to these inquiries, I carefully read the relevant patents (U.S. 5,450,490, 5,764,763, 5,581,800, and 5,787,334) issued to Arbitron. In addition to the patents themselves, the file folders ("file wrappers") maintained by the U.S. Patent and Trademark Office (USPTO) containing correspondence between the inventors and the patent examiner during the application process are in the public domain and available to any interested party.

Since most of the relevant patents were filed and issued in the late 1990s, it is possible, and perhaps likely, that Arbitron has advanced the PPM technology beyond that described in its patents. Nevertheless, the fundamental basis of PPM as described in these patents is entirely consistent with Arbitron's public literature.

The PPM system is just one example of "watermarking" technology, which has been subject to extensive research for the past two decades. The Arbitron PPM system embeds watermarks with station identification codes into the audio program at the time of broadcast using an encoder in each individual radio station's transmission chain. Portable PPM decoders then identify which stations the wearers of these "people meters" are listening to.

In order to embed the digital bits that make up the identification code, watermarking modifies the original audio by adding new content or changing existing audio components. The goal of an ideal audio watermarking system is to be 100% reliable in terms of embedding and extracting the watermarking data in all "typical" listener scenarios while remaining 100% inaudible for all "typical" program material. These goals underscore a paradox: 100% encoding reliability requires audible watermarks. Conversely, to achieve total inaudibility, watermarks cannot be present on some material, which sacrifices reliability. Anecdotal reports from radio broadcasters say that Arbitron lowered the watermarking energy in response to complaints about the watermarking being audible in certain circumstances. Trade offs must always be made in audio watermarking systems to balance audibility and reliability.

Some radio broadcasters question whether the PPM watermarking system is as reliable as Arbitron claims. From my analysis, I conclude that the answer is both *yes and no*, depending on the definition of typical program material and typical listening environments.

Real-world conditions may not always match the assumptions made by the PPM system designers. This article explores potential consequences of station programming and

25-Seven Systems, Inc.

1085 Commonwealth Avenue #360, Boston, MA 02215 USA
+1 (617) 789-4673 • e-mail: info@25-seven.com • <http://www.25-seven.com>
© Copyright 2009 25-Seven Systems, Inc. All rights reserved

listener behavior that deviate from the PPM system designers' assumptions of typical. After conducting a modest literature search, I found no published papers that describe the behavior of the PPM system in real-world conditions. Because of the absence of empirical data, this article also considers how an empirical evaluation might be conducted.

Theoretical Foundation

To one degree or another, all watermarking technologies use the well-known perceptual principle of "masking," which was first reported in the early 20th century and is a core technical basis for mp3, AAC, and a host of data-rate reduction schemes. In simple language, a loud burst of energy at one frequency will deafen the human auditory system to certain other audio components at nearby frequencies for a period of time before, during, and after the loud signal.

Consider the following illustration: A tone burst at 1.1 kHz with an intensity of 0 dB will hide (make imperceptible) an added signal at 1.11 kHz with a level of -30 dB for a period of about 10 ms before the burst and as much as 50 ms after the burst. However, modern signal-processing techniques can still detect the existence of this added 1.11 kHz component even though the ear cannot. This is the basis of PPM and other similar watermarking technologies that use masking for determining the frequencies and intensity of the data that can be added for the station-identifying watermark.

The PPM system constructs 10 spectral channels in the region from 1.0 kHz to 3.0 kHz. The original program audio energy in each channel is evaluated for its ability to mask an added component. If that masking energy is insufficient, nothing is added. Conversely, if the energy in a channel is large enough, a tone is injected, chosen from one of four possible frequencies within the channel. For example, the channel centered at 1058 Hz might have one of the following four frequencies injected: 1046, 1054, 1062, or 1070 Hz.

Each of the four frequencies represents 2 bits of information. If we assume that this process repeats at a 500 ms rate, using all channels provides 40 bits per second or 2400 bits per minute of watermark code. Let's further assume that a radio station is credited for a listener if any code is correctly detected within a 3-minute interval. With the very large number of encoded bits generated in 3 minutes (2400 x 3 = 7200 bits) and a station's identification data needing perhaps only 50 bits, there is massive excess capacity for redundancy, error correction, and for audio that does not have enough high-frequency content for masking.

The system should work perfectly. But nobody, including this author and Arbitron, expects the PPM system to be perfect. Every technology has limits, which should be understood by users of a technology.

25-Seven Systems, Inc.

1085 Commonwealth Avenue #360, Boston, MA 02215 USA
+1 (617) 789-4673 • e-mail: info@25-seven.com • <http://www.25-seven.com>
© Copyright 2009 25-Seven Systems, Inc. All rights reserved

Hidden Assumptions—Encoding

While “most” audio, especially conventional music, produces a full band of energy above 1 kHz, some audio has little energy above 1 kHz and some has virtually none. Obviously, silence is the extreme case, where there is no possible masking for inaudible watermarking.

When audio produces only a modest amount of energy above 1 kHz, only a few channels will be available for carrying *strong* watermarking information. In such cases, the PPM system can encode watermarking only at very low levels, which makes it fragile when being received.

Examining the spectrograms of certain instruments sparse in musical overtones can provide predictive information on when the PPM will be stressed. Considering that a male fundamental pitch might be as low as 80 Hz, some announcers may have a speaking style that is weak in high frequencies. Depending on the structure of the vocal cords and articulation style, there may or may not be any energy at the 12th harmonic of that pitch (which happens to be the center frequency of the first channel of the PPM encoder) for some announcer.

Because fricative phonemes (such as /s/, /z/, /th/, and /f/) contain a broadband hissing component that is like white noise, they can encode large amounts of data. But some announcers may have weak or rapid articulation of such fricatives. Consonants, although short in duration, are good for PPM; pregnant pauses and halting delivery are not. Speaking style matters.

While the typical radio program may produce perfect watermarking performance, and while the average reliability over the universe might be 99%, there are likely to be some announcer voices, vocal articulation styles, and specific genres of music that belong to the 1% failure cases. If a particular program on a particular station is one of the failure cases, that program might experience “bad luck” in its audience ratings.

Hidden Assumptions—Decoding

Even though the PPM system may successfully encode sufficient bits into a radio program, the portable decoder may or may not be able to detect those bits if the listening environment deviates from the designer’s assumptions. Even if the station’s confidence decode monitor, operating with a wired connection to the encoder or a monitor tuner, confirms that there are sufficient bits embedded in the transmitted signal, there is no assurance that the portable monitor can extract those bits.

Specifically, the usable signal-to-noise ratio between 1.0 and 3.0 kHz varies dramatically depending on the environment. Not only does the power in the watermarking depend on encoding channels and playback volume, but a listening environment may also be very quiet or extremely noisy. A hostile listening environment is less likely to have deleterious

25-Seven Systems, Inc.

effects when all 10 spectral channels have watermarking data, as some channels will get through. But if only a few channels are encoded, consistent background noise may overpower them.

In a highly absorbing listening space with plush carpets and padded chairs, high frequencies are more readily attenuated than low frequencies. Some listeners may use a radio broadcast as background and not care that they are only hearing the low frequencies. Depending on the loudspeaker in the radio, high frequencies may project in a highly directional narrow beam while the low frequencies are omnidirectional. High-frequency content may also be blocked and absorbed if the listener has attached the portable PPM monitor to his or her clothing in a way that it faces the upholstery of a plush chair or if the listener has placed the device in a purse or backpack.

Given the wide range of listening environments, ranging from always decoding to never decoding, a radio station can influence the PPM performance in the intermediate cases. Marginal cases will decode correctly if the watermarking energy is made stronger, and this factor is influenced by the properties of the audio source. Even though radio stations can only influence encoding, they should keep in mind the full system of encode and decode processes.

Hidden Assumption—The Influence of Randomness

Statistical measures of reliability have yet another hidden assumption: randomness.

Even if the PPM system were only 90% reliable, the system would be acceptable if the errors were distributed uniformly across all programs and listeners. Errors would average out. However, a system that was an impressive 99% reliable over all announcers, program types, and listening devices might have the failures concentrated in a particular intersection of programming type and listener subculture. The fallout from such a failure might be further amplified by an inadequate audience sample size.

Statistical assertions about accuracy almost always assume that bell-curve models represent the real world. But very few real-world situations actually fit the bell-curve model. If the errors are not random, there is a problem. For an in-depth treatment of this kind of statistical failure, see [Nassim Taleb's](#) book *The Black Swan: The Impact of the Highly Improbable*.

Avoiding “Gaming” of the System

In the absence of empirical data about the frequency and severity of failure scenarios in the real world, we can only speculate about their existence. Nevertheless, I conclude that there will be some cases where the PPM system fails simply because the program content and listener environment do not match the assumptions made by the designers of the PPM system.

25-Seven Systems, Inc.

1085 Commonwealth Avenue #360, Boston, MA 02215 USA
+1 (617) 789-4673 • e-mail: info@25-seven.com • <http://www.25-seven.com>
© Copyright 2009 25-Seven Systems, Inc. All rights reserved

A radio station has no direct way of modifying how portable decoders behave, nor should it. Gaming the system at the listening end would be no different from stuffing the ballot box or falsifying paper diaries. It is in the industry's best interest that a firewall be maintained between stations and individual participants in the audience measuring process.

Nevertheless, insuring that audio leaving the station is optimized for successful PPM encoding is every station's business. A program director who understands which audio makes the PPM system more reliable has an advantage over a competitor who does not. Knowledge that is not public has high monetary value, not unlike that used for insider trading in the stock market.

To avoid gaming the ratings, I believe all broadcast stations should have a deep understanding of how the PPM system works in the real world, not in the ideal world of the designers' laboratory. Stations should understand why their confidence decoders are kicking out failure indications.

Testing Your Radio Station

A radio station can test its programming as well as explore the listening assumptions in the world of PPM. There are two basic approaches to evaluating your environment.

First, the input to the PPM encoder can be subtracted from the output of the encoder to create a "mix-minus" of the encoding. The subtraction removes most of the original program and allows the channels to be seen clearly with a spectrogram analysis of the difference. This will indicate which channels are being used to encode data and measure the total power in the watermarking information. Given this measurement technique, every program can be evaluated for its watermarking robustness, showing the ability of the PPM system to carry watermarking information on specific audio samples without examining the encoding itself.

Moreover, the Arbitron decode confidence monitor provides an output that indicates if watermarking decoding was successful. If the encoding process is marginal because of weak channel encoding, then it may take longer for the decoder to find the embedded data, if it finds it at all. A long response time to acquire the station ID watermark serves as a warning that a non-ideal listening environment may result in a failure to detect the ID.

Second, the Arbitron confidence decoder monitor can be placed in a listening setting deemed by the station as typical of its real-world listeners. For a drive-time talk show, the environment might be a truck or automobile with the windows rolled down. For a late-night show, it might be a bedroom. The microphone used by the decoder device as its input source may be at some distance from a radio loudspeaker and there may be "environmental" obstacles, such as a chair blocking the high-frequency channels but not the low frequencies. Effectively, one can determine the required signal-to-noise ratio in the

25-Seven Systems, Inc.

1085 Commonwealth Avenue #360, Boston, MA 02215 USA
+1 (617) 789-4673 • e-mail: info@25-seven.com • <http://www.25-seven.com>
© Copyright 2009 25-Seven Systems, Inc. All rights reserved

listening environment for successfully decoding a specific received program for that program's target audience.

Keep in mind yet another hidden assumption: There is no evidence that the performance of a station's confidence decoder monitor is the same as that of the portable PPM carried by listeners. The confidence monitor is fed by a direct-wired connection, has a large power budget, and is seemingly without constraints on the computational power dedicated to decoding. In contrast, the portable decoder may have severe constraints related to power consumption, volume, and computational capabilities. On the other hand, the portable version may have front-end analog processing and filtering to optimize the extraction of the acoustic signal. Obviously, the portable device is a better test vehicle since it mimics the real world.

I believe that the following general principle is valid: The PPM system is designed to handle specific difficult cases, but a combination of difficult cases is likely to produce failure. A radio station must therefore determine if such combinations exist for any of its programming.

While acknowledging Arbitron's assertions about reliability and respecting its proprietary intellectual property, any radio station can run experiments to measure performance and robustness. The results of such experiments will show whether certain programs in certain listening environments break the technical assumptions upon which the system operates.

Looking Forward

This article was written to frame questions and to explore what is not currently known. When there is more data, the implications will become clearer. At this point, my examination of the PPM system is purely speculative, based on patent disclosures and anecdotal reports distributed in radio industry trade publications. That the various sources of information are consistent does not prove that the picture is accurate or complete.

My best guesses at the scenarios most likely to fail are:

- Audio program material with low-level high-frequency content, implying low-level watermarking tones.
- Listening in a noisy environment or with the monitor positioned to receive only a muffled signal, increasing the likelihood that the PPM monitor will not correctly detect the stations' ID.

When combined, these individual weaknesses might create a "perfect storm."

Making the audience measuring process more transparent and reliable, as well as providing an unbiased playing field for all participants, will benefit all the stakeholders in our industry. Adopters of any innovation owe it to themselves to understand how the technology works. Technology is science, not magic, and providers of technological solutions are best served by an informed and enfranchised user base.

25-Seven Systems, Inc.

1085 Commonwealth Avenue #360, Boston, MA 02215 USA
+1 (617) 789-4673 • e-mail: info@25-seven.com • <http://www.25-seven.com>
© Copyright 2009 25-Seven Systems, Inc. All rights reserved

About the author

Former MIT professor Dr. Barry Blesser has amassed a reputation for technologies ranging from digital audio workstations to handwriting recognition. A recognized authority on audio technology, both as an inventor and an expert witness in several high profile technology cases, Blesser has designed radio broadcast and professional audio products for a wide range of companies, including EMT, Orban, and Lexicon, and served as chief technology officer for Studer. For the past four years, he has written a regular column for *Radio World Engineering Extra*. Blesser is a past president of the Audio Engineering Society, a consulting technical editor of its journal, and a recipient of its Silver and Governors medals. MIT Press recently published his book *Spaces Speak, Are You Listening? Experiencing Aural Architecture*.

25-Seven Systems, Inc.

1085 Commonwealth Avenue #360, Boston, MA 02215 USA
+1 (617) 789-4673 • e-mail: info@25-seven.com • <http://www.25-seven.com>
© Copyright 2009 25-Seven Systems, Inc. All rights reserved