The Human Wonderland of System Complexity

By Barry Blesser @2010
www.blesser.net

In the 1970s, a few of us realized that signal processing could be used in commercial products for the audio and broadcasting industries. We had the knowledge and skills to be successful, and the world viewed us as technical wizards, awarding us high stature and monetary compensation. My highpoint was the development of an electronic concert hall in a small box (the EMT-250), thereby replacing large and inflexible reverberation chambers. We rode the wave.

Almost half century later, products that use digital signal processing are ubiquitous, inexpensive, and magically powerful. The high art of yesterday's engineers is now the accumulated wisdom of today, repackaged in inexpensive chipsets and enabling thousands of commodity products. Sophisticated technology is now a commodity. For an engineer, system design replaces component design.

So what do we mean now when we describe something as a "system?" A system is a collection of elements that interact with each other such that its properties cannot be found in any of the individual pieces. Now more than ever, technology is just one of the elements in broadcast systems that also include many different types of people: investors, managers, listeners, colleagues, advertisers, competitors, journalists, and of course, engineers. To survive, engineers must have, or acquire, the necessary skills to design and maintain systems that include the behavior of people.

How then can an engineer in the 21st century regain the stature and compensation of his earlier counterpart in the 1970s? The answer is to redefine the scope of a system such that people are now included as a major element. Human behavior and psychology follow well defined rules but they are definitely not the same rules as those for hardware and software.

AUDIENCE MEASUREMENT IS A SYSTEM

Over the decades, new elements have been continuously added to systems. The 1960's hardware engineers had to retool their skills in the 1980's with the advent of software. Electronic processing was added to mechanical systems; user-interface design was added to computer systems; digital signal processing was added to audio systems; and currently people have been added to almost every type of system. The idea of thinking about people as being *in* the system is relatively new.

People have properties that never become obsolete, unlike specific technologies. Evolution works on a very long time scale, and our biological brain has not changed much over the centuries. In contrast, my technical skills of the 1960s became obsolete in the 1970s, and my skills of the 1970s became obsolete in the 1980s. But my understanding of human psychology, which I have been acquiring for the last 40 years, continues to become more sophisticated as I

gain more experience. Wouldn't you like to have some skills that, like fine wines, got better with age?

While there are many examples of this new kind of system in the 21st century, I will focus on one particular broadcast system as an illustration. It also happens to be a hot topic. Audience measuring technology provides a perfect illustration of how people determine the personality of the system.

When the audience measuring system in the radio industry used paper diaries, our industry was using the technology of our ancestors: paper and pencil. Since there was only one technical solution everyone accepted its properties, good or bad.

Beginning in the 1980s, research scientists began to consider using signal processing technology to embed digital codes into an audio signal without producing any audible degradation. These codes (watermarks) could be used by a broadcast station to label audio as having originated from their station. Monitoring equipment carried by a listener then detects the presence of these codes, thereby connecting the listener to the station. Sophisticated technology could now replace primitive paper.

Arbitron and many other companies considered the problem of watermarking to be a difficult technical challenge. Eventually, after a few decades of research and development, engineers created a system that worked well in laboratories and pilot studies. After successfully testing and adjusting the system, Arbitron has been replacing the paper diary method with their new automated electronic solution, the PPM (Portable People Meter) in an increasing number of markets. From a technical perspective, the design is very elegant.

During the first half of 2009, executives at our company (25-Seven Systems) were approached by friends and colleagues to help investigate why the shift from paper diaries to the PPM system had adversely affected particular radio shows. Evidence was building that something was not right. There were suspicions that there might be a technical explanation for changes in ratings.

TECHNICAL VS. HUMAN PROPERTIES

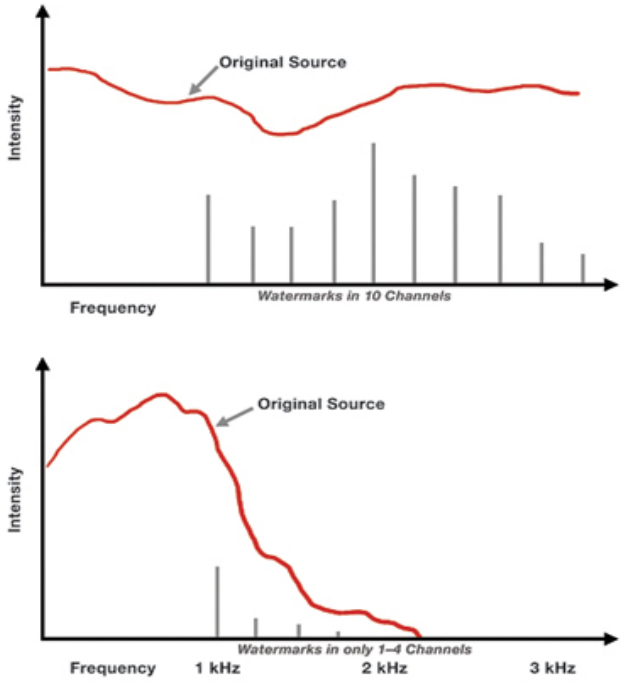
As inquiries continued to arrive at our main office, I decided to examine PPM as a complex system. After analyzing Arbitron's patents and talking to many knowledgeable people who had experience with PPM, I wrote a white paper, "Technical Properties of Arbitron's PPM System," which summarized everything that I understood about the PPM systems at that time. I wanted to share my understanding of how the PPM system might actually behave in the real world of imperfect people listening to a wide variety of audio. If you have not read the paper, you can download a copy using the link on our website page: www.25-seven.com/blesser.html.

I was not the only one interested in the implications of PPM. In November 2009, I was invited to speak with the Congressional Committee on Oversight and Government Reform in Washington. This committee was researching the PPM system in preparation for a congressional hearing on the potential impact of PPM on minority stations. While I had been

primarily focused on the technical properties of PPM, the committee was concerned with an entirely different side of the story: people. It is well worth your while to view at least some of the hearing using another link on the same page of our website.

The designers of PPM, while focusing on technical challenges, appear to have underestimated the relevance of the "people" part of the system: listeners in specific sub-cultures with unique attitudes, values, and life styles. People are not electrons with identical properties; they are all different. Furthermore, the behavior of listeners in some cultural groups can collide with the inflexible technical properties of PPM to create unanticipated side-effects that degrade the reliability of the system. Apparently, the designers of the system had not realized that the skills of psychologists, sociologists and anthropologists were also needed as part of the system design process.

Before we explore how technology and people together create a complex system, I need to explain a few basic aspects of the PPM encode-decode algorithm. Watermarking energy, which contains the digital codes identifying a station, is injected in the spectral region from 1 to 3 kHz at a level well below the audio energy in this band. Because of the psychoacoustic property of masking, these watermarks are inaudible to the ear if their energy is significantly lower than the corresponding energy in the original source. Sounds that have strong energy in this spectral band can mask strong watermarking signal strength. Sounds with weak energy can only mask weak watermark energy, and faint sounds with little or no energy cannot mask any watermarking energy. Watermarking energy depends on the source audio.



Top: High watermarking energy when the audio is full energy over the spectrum
Bottom: Weak or missing watermarks for audio that is mostly low frequencies.

There are many examples of how technology and people interact in unanticipated ways. Consider the real life situation of a family living in an apartment with open or acoustically porous doors. Assume that a teenaged son is listening to a hard-rock station in his bedroom, and that his father, wearing the PPM device, is listening to a talk-radio program in the living room. The PPM device hears two watermarks from two programs, and it will decode the stronger one. But watermarking energy at the PPM device is determined by two factors: (a) proximity between the radio and the device, and (b), the amount of watermarking energy in the source. The watermark energy for the son's program might be 40 dB stronger than that of the father's program, which is enough to overcome the extra attenuation of a longer distance to the device.

In such an environment, the talk-radio program does not get credit for having a listener in this household. And if both the son and father were independently participating in the Arbitron listener panel, the hard rock program gets credit for two listeners.

UNEXPECTED SYTEMATIC BIASES

Another example of unintended consequences arises from the assumption that any system will produce random errors, and such errors will average out. In contrast to random errors, systematic errors become strong biases in audience ratings. Consider an example of a systematic bias.

For fashion obsessed youth and young adults, wearing the PPM dongle would not be "cool" today as it might have been in the 1990s when a pager was a status symbol. On the one hand, a cell phone evokes connectedness, social relevance, and can be customized to individual tastes with apps, skins and ringtones. On the other hand, the larger PPM dongle has become a symbol of an electronic dog collar worn by people tied to a job.

Attitudes towards devices depend on social status, and those with a cool image are likely to refuse to participate in Arbitron's audience panels if it means carrying around a visible dongle for 8 hours a day week in and week out. Had the PPM technology been incorporated into a cell phone, the story might have been different.

Those who are suspicious of governments and organizations might have been willing to fill out a paper diary for a couple weeks, but may fear that the dongle is an electronic tracking device. Dongle symbolism can result in under-counting of some cultural groups. A pencil and paper diary gives participants full control of what is happening, while a dongle demands hidden (mysterious) interaction with a command center. This may evoke fears that "big brother is watching" for a segment of the population intent on staying hidden. Even to a knowledgeable engineer, the unverifiable nature of a sealed dongle becomes a psychological problem. Knowing what I know about technology, computer viruses, corporate machinations and newspaper reports of unethical behavior, do I really want to trust a corporate spokesman who tells me to blindly trust them?

Using myself as an example with my life style and value system, I would never participate in any audience measuring activity that required me to wear a dongle week in and week out for

months or years. But I might have been willing to fill out a paper diary in the evening for a few months. I, and others like me, represent a sub-culture that remain inaccessible to the new audience measurement process. But we might be a critically important group to advertisers.

PERVERSE CONSEQUENCES

To combat the resistance of people to participate in audience panels, Arbitron offers a modest payment per month. A minimal payment produces a motivational incentive for those lower income people in desperate need of additional cash. Those with wealth, who may be an important segment for advertisers, are very unlikely to be motivated by an additional $10 per month.

From a purely economic perspective, even though the PPM can be fully automated it is now proving to be expensive to train and retain panel participants. The additional cost of managing participants pushes more cost on to the broadcast industry. Automated technology can actually have a high human maintenance cost.

As a final example of unintended consequences, consider that many radio stations discovered that specific programs have sounds that only weakly encode the watermarking, namely certain types of jazz and male talking heads. Because radio stations have no way to influence the encoding process, their only choice is to modify their programming to better match the design assumptions made by PPM developers more than 15 years ago. Programming decisions are now driven by the technical properties of PPM not by listener preferences. Without realizing it, designers in 1994 sitting in their laboratory were actually changing listening options in the 21st century and perhaps for decades.

Combining these issues leads us to the conclusion that audience ratings may have a bias towards those who are poor, uninterested in being cool, and enjoy full bandwidth music. Biases are systemic errors unique to particular scenarios, sub-cultures, and broadcast preferences.

In contrast, random errors assume sampling a uniform population with a large sample size. Human populations are seldom uniform. In fact, averaging as a statistical technique assumes that all errors arise only from additive random noise. An engineer without a background in population sampling might not realize that a measure from a particular population in a particular scenario does not predict performance with other scenarios.

REAL WORLD BETA TESTING

Notice that these systemic biases all arose from how people interact with the technical properties of the PPM design: packaging symbolism and varied watermarking energy. I have no doubt that the design would have turned out differently if the developers had had the foresight to consider what they didn’t know, and to recognize that a complex system, which includes human behavior across multiple subcultures and shifting social norms, cannot be solved only with technical wizardry.

Are my descriptions of biases real or just hypothetical possibilities? What is truth in audience measurement? Unfortunately, a methodology to establish truth, even for one week in one city, is pragmatically impossible. However, because there is data for two methodologies, PPM and paper diary, and to the extent that they differ, one or both must be in error. If one reports an audience size of 10,000 and the other reports 20,000, then at least one of the two must be off by at least 5,000 listeners. In other words, the existence of a second method automatically highlights biases that produce winners and losers. Each station will prefer that system which provides the most favorable ratings. The difference in ratings now becomes political and economic, not technical. The congressional inquiry arose because a significant number of losers were minority owned stations.

These issues have come to light only after PPM was deployed on a wider scale; designers call this a Beta-test. Until a design has been exposed to a large number of people in real world, one will never know what scenarios will prove to be relevant. When Arbitron says that the system was extensively tested, they are correct, but controlled tests with specific scenarios do not represent the world as it is. It is impossible to discover all of the scenarios by thinking.

KEEP PEOPLE IN THE EQUATION

Given the complexity of the issues discussed, Arbitron has taken the position of using secrecy, rather than an open discussion, as a way of responding to challenges to their system. In many private conversations, I have been told that secrecy is required to prevent stations from "gaming" the system.

Paradoxically, secrecy actually creates a situation where a very smart advertiser could successfully take advantage of defects in the system. Given that ratings have biases, with the appropriate research an advertiser could estimate the size and nature of the bias for a particular program. For example, the reported audience size might be 10,000 but by including the bias, the actual size might be estimated as being 15,000. The advertiser pays for air time on the smaller measure and gets 5,000 extra listeners.

Audience ratings are a current topic, but in the context of this discussion PPM is just one example of how 21st century systems are now dominated by the properties of humans. Applications where human beings and technology fuse together to become a single system are distributed throughout our daily life. The user interface of your favorite word processor is a connection that creates a system with you in it. Similarly, executives participating in a remote audio-video conference are also a system with people and technology.

Engineers who are comfortable with technology need additional skills besides their technical wizardry. People are part of the system, and people are not black-boxes that follow the rules of hardware or software. People are the ultimate challenge for engineers who design, deploy and maintain systems. And finally, whatever skills you acquire in understanding people, be they listeners or executives, those skills will never become obsolete. Ignore people at your peril.