

## Communication Theory and Argent™

At the most basic level Argent is concerned with embedding an informational message, considered abstractly as signals, into existing signals in a manner that is secure from unauthorized reading and adaptable in handling changes introduced to the carrier signal and its encoding.

Argent deals with basic read-out or access security by using pseudorandom locating methods to determine where in the carrier signal a message is embedded. In situations where the carrier signal is sufficiently different from an informational signal, the security is provided by making a trial-and-error search for decoding impractical. Location is a logical concept that can include both temporal position or frequency position though not necessarily limited to these options. In particular, the Argent model includes various combinations of locational modalities. If the informational message is considered on a bit-by-bit basis then location determines where each particular bit is stored in the carrier signal; however, the model is not limited to a single bit per single location implementation. In particular, it includes a "single bit to multiple location" implementation. This implementation allows for additional robustness and also for the encoding of multiple message bits to abstract signal features such as frequency or amplitude curves, and other statistical signal measurements, which may be changed without affecting perception of the carrier signal.

The robustness of an Argent encoded message can be dealt with in what is considered "the mathematical theory of communications". An important paper on the subject, "The Mathematical Theory of Communications", was published by Claude E. Shannon and Warren Weaver. The paper is well-known to those skilled in the art of communications theory, particularly as applied to computers and digital systems. It includes some basic theories and formulations concerning transmission rate measured in bits or symbols per second, the effects of noise on transmission, methods to counter the effects of said transmission, and methods for making particular transmissions as robust as possible. This latter subject will be examined strictly in terms of what is called error detection and correction or "error coding" as a whole.

The field of error coding is quite important in any digital transmission system. In principle, error coding involves modifying the symbols comprising the information to be transmitted in such a way as to make transmission and subsequent reception as unambiguous as possible in the presence of given existing noise. Coding is managed in such a way that, provided the error rate in the transmission does not exceed a certain probability, the full message may be reliably recovered to an arbitrary degree of certainty. Typically this is accomplished by creating redundant

the transmission. Note that it is never entirely possible to eliminate or be 100% percent sure that what is received is what was sent. Error coding allows one to minimize the percentage uncertainty, equivocation, to an arbitrarily insignificant amount. This allows application of information transmission. So, transmission accomplished where the chance of a mistake is, say, 1 in 3 trillion, proof of accuracy is required, an analogy with DNA evidence in a court case becomes a relevant benchmark.

An important reference point in error coding is the error rate of 50%, which represents situations where each bit transmitted has an independent probability of reception in error of 50%. Put another way, what is truly subject to random chance. The message may be right, or wrong, on a completely random basis. The Shannon paper makes the point that this is the theoretical limit of error coding systems. A transmission rate of 50% is impossible to recover from, since it would require redundancy in the encoding process for subsequent recovery. A rate below 50% can theoretically be coded. Generally, there is a positive relationship between error rate survivability and required redundancy, peaking with an infinite redundancy requirement at 50%. Note that a rate of higher than 50% would theoretically allow a reversal of the code to create an error rate of  $100\% - R$ .

An important point is made in the Shannon paper regarding transmission of information in the presence of truly random, independent (50%) noise. Although one might compute a measure of transmission of bits or characters per second under this circumstance, it is false, for the simple reason that one can never be sure what was received is what was transmitted, so the effective transmission in this situation is 0.

Because of this fact, and because of aspects of typical applications of embedded signaling systems, as those described by BBN, CRL, and to some extent Digimarc and HighWater Design, are derivatives of basic modulation systems which modulate information, one way or another into a signal. The expected result is to have some variation from true transmission: these embedded signaling systems adhere to certain principles. A good example is the modem, which modulates an analog carrier with digital data, and expects that analog noise will cause the carrier to vary continuously for each transmission. The carrier signal is not expected to be exactly the same, twice. Because of this, an embedded signaling system built on such a model typically attempts to embed a signal which carries its own, separate from the carrier signal. In reality, these systems are called "composite signal" embedded signaling, because, in effect, they combine two complete, independent signals into one. This is because of the "50% problem" described above. The embedded signal can very well convey information in the absence of the carrier signal.

important requirement for typical transmission systems. Such a system is called "embedded signaling" because it attempts to structure the signal in such a way as to mask it using characteristics of the carrier. For instance, if the carrier signal is audio, which is meant to be heard, the embedded signal is constructed so that characteristics of the carrier prevent the embedded signal from being "heard". Nonetheless, it is there.

Argent, on the other hand, is intended for a very specific application and does not necessarily dictate that the carrier signal should vary from one playback to another. Since Argent was conceived to protect copyrights and "rights," and because a copyrighted work is generally a fixed signal, it allows for an interesting opportunity. The carrier signal, which may be analog but not necessarily, digital audio or video, should be exactly the same from one playback to another. Even if it is not exactly the same, it should be very close to the original. Any variation represents an error in the carrier signal, whether intentional or not. In typical modulation systems the carrier signal means to transmit the informational signal, and is of no other use. Circumstances are quite different with copyrighted signals such as audio or video. Such errors are imperfections in the signal, and the more of these errors there are, the more the quality of reproduction of the original is adversely affected. In the field of copyrighted material, a nice feature of Argent is that this tends to reduce the commercial value of the copyrighted signal.

It is our belief that those who promote traditional carrier schemes for embedded signaling purposes in order to protect copyrighted works, in essence, cannot see the forest from the trees. They are making an implicit assumption that the carrier signal does not matter, when in fact it is all that matters.

Because Argent recognizes the value of the carrier signal, specifically the unique situation that, in principle, the carrier signal is invariant from one playback to another, it can take advantage of this invariance with respect to mathematical communication theory. In particular, Argent is not bound by the restraint that the carrier signal can vary arbitrarily from one playback to another, and can therefore exploit what we term "free bandwidth", or the ability to transmit information by creating an additional signal which on average transmits only half the information in the original. Work by Craven and the late Gerzon provides valuable theoretical parameters for audio "free bandwidth". Substantiated within the application, considerably fewer (one half on average) changes in the carrier signal in terms of modulation, because the other half of the information can, in effect, be transmitted by the original carrier. This fact would tend to increase the transmission rate, while decreasing the effect on the carrier signal versus other systems, everything else being equal. An Argent "embedded signal" is not independent, it depends on the carrier.

signal itself, and on average, it depends on the carrier signal information. Thus, widespread use of the term "digital watermark" is a mistaken definition for systems other than Argent. The more common definitions, "signature" or "fingerprint", which are used for unsecured authentication (that is, for copywritten multimedia signal ownership, unlike the design goal of Argent, are more appropriate. Note that in a traditional transmission, if the carrier signal is uncorrelated to the information signal, there is a corresponding error rate condition. But because Argent was designed specifically to work in a situation where the carrier signal is invariant from one reproduction to the next, this "error" rate is in fact, not an error rate at all.

Theory aside, what happens when the carrier signal does vary on reproduction? Specifically, what if the carrier signal is analog, which has inherent noise? What if errors have intentionally been introduced with the intent of erasing an Argent watermark? Recall that the only limitation on the error rate is 50%, or truly random. What makes Argent work in these situations is that because the inherent nature of copywritten carrier signals is such that they must at least be arbitrarily close to the original "commercial" value, the error rate caused by variations in the carrier signal is not likely to affect an Argent encoding with the full 50% error condition. The illustration is in order.

Suppose we seek to encode an information signal into a copywritten carrier signal, and that the information signal is entirely uncorrelated to the carrier signal. This means that on average, if compared bit-by-bit, the information signal to the Argent signal, then one half of the bits will match by random chance. This example assumes a very simple approach to encoding where 1 bit of the information corresponds to 1 bit of the carrier signal. While a secure and practical Argent implementation would surely be more complex than this simple, the principles of the example hold. So, the message bitstream which depends on the original carrier signal for one half of its bits and on changes introduced to the carrier signal by the Argent encoding process, for the other half of its bits. Now assume that after encoding, the composite (encoded) signal is changed. If the mapping which encodes the Argent message is one-to-one, that is, every bit of the carrier signal carries Argent information, then in order to induce enough changes to make an Argent message theoretically unrecoverable, the composite signal must have to be completely randomized. That is, for every bit of the carrier signal, one would flip a coin, and if that coin came up, say, heads, one would flip the bit of the carrier signal. Not only would the Argent message be unrecoverable under this circumstance, but the composite carrier signal would be worthless noise.

However, this example differs from a secure implementation of Argent which minimizes degradation of the carrier signal. So, let us make

assumption. Assume the mapping for Argent encoding is that for every sample in a digital audio recording, Argent will cause a change to one of every 4 samples, or a 25% density. If a sample is 16 bits, then the density is really 1 in 64 bits. In theory, to make this message unrecoverable, one would have to cause 50% of the Argent message bits to change, which would normally correspond to 1 in 128 of the composite signal bits. This would have a much lesser impact on the composite signal. However, there is an issue here. The 1 in 128 figure assumes that one knows exactly which composite signal bits are actually Argent message bits. If one bit randomly changes to 1 out of every 128 composite signal bits, then the independent hit rate would be 1 in 128, assuming the agent of change has no information regarding which bits are Argent message bits. So, to induce a level of errors in the Argent message to make it then unrecoverable, given one has no information which bits of each 128 bits are Argent information, the independent hit rate must be 1 in 2, which means that the agent must change, on average 64 of every 128 bits, or 50% of the composite signal.

There are a lot of qualifications in practical use. First, in a composite to one bit of sample encoding scheme, one could make some simple assumptions regarding which sample bits represent legitimate candidate bits for Argent encoding, and so reduce the independent hit rate proportionally to the requirements, and hence, total change requirements. In addition, this discussion ignores the redundancy requirements, and does not illustrate the requirements when the error coding rate accounts for rates significantly less than 50%, as must be the case, since it is impossible to reach the required recovery. So, exactly what a change agent must do will depend on the error coding rate of the Argent encoding. This in turn, impacts how the Argent message must be, and how many copies may be placed in a composite signal.

However, what should become clear is that Argent accounts for the difference between copyrighted carrier signals, and simple modulated signals, and takes advantage of this difference to provide "free" flexibility, which in turn means Argent can be configured to make fewer changes to the carrier signal, while effecting higher bit rates than other systems where the carrier being equal. In addition, Argent allows rights holders to make determinations in configuring exactly how encoding is accomplished, which allows the tailoring of encoding to be resistant to arbitrary levels of error. Such tailoring is not without its tradeoffs. However, we feel it is important to provide flexibility as an option. We note that assuming carrier signals of arbitrarily long length with respect to a shorter Argent message, the user is free to choose a sufficiently high level of error coding to provide significant robustness in the encoding. The only question is how much of the carrier signal is required to carry a single watermark.

and how many copies of such a message can be embedded in the c signal.

Argent could very well take advantage of such masking techniques as known in the art of multimedia engineering to further reduce its signal quality as well as limit its encoding parameters to su translations as MPEG compression and analog conversion. Again the tradeoffs which sacrifice some of the encoding bandwidth versus r The point is that the Argent model is built on an optimal fr regarding the protection of rights associated with signals of valu standpoint, one would expect to have the most operating headroo which to achieve optimal results.

Finally, we have discussed certain probabilities and made the po theory, Argent may independently encode only one half a complete Even a naive reader may ask, what good is half a message, and how that the full message I got out, is the one that went in?

First, we remind you that although Argent may induce changes repr one half a message, the rest of the message is in fact carried l signal itself. The rest of the answer, once again lies in probab that one encodes a message of an arbitrary length, say N bits. Th probability of finding a random match to this bitstream, which m every bit, is  $1/2$  to the Nth power. So, the longer the message, probability of a false positive. In addition, Argent takes some o to ensure trust in its results. The software employs special minimum bit length to delineate a valid watermark. This esta minimum degree of surety. In addition, the extra bandwidth prov Argent can be used for such things as digital signatures and checks provide their own independent degrees of verification. So, as a f question of false positives, very evident in typical "embedded systems is akin to asking, "what if the monkeys DO type Shakespea answer is simply that the Argent user can predetermine the probabal happening, and set it to an acceptably low level.