### **WIRELESS APPLICATIONS OF ARTIFICIAL INTELLIGENCE**

*AI Finds Better Solutions* 

David E. Newman and R. Kemp Massengill

#### Executive Summary

 6G is going to be complex and fast-paced. Demands for higher throughput are growing exponentially. Our ambitious 6G goals clearly cannot be achieved with humans in the loop. Only AI can make 6G truly successful. Three wireless applications are outlined below that provide enhanced communication performance with AI support: (a) efficiently identifying and correcting message faults without a retransmission, (b) adjusting beam parameters in real-time without beam scanning, and (c) selecting an optimal modulation scheme based on current message fault rates. It is difficult to imagine 6G without these AI-based solutions.

 The innovations described below are not aimed at standards; they represent business opportunities. Companies implementing these methods will obtain a competitive advantage. The innovations disclosed below will enable wireless companies to provide better services and better performance, increased customer satisfaction, growing sales volume, and higher profits for the provider.

#### Artificial Intelligence Models

 AI excels at complex, non-linear, multi-variable problems requiring instant, "good-enough" decisions, in a rapidly evolving environment - such as 6G. Training is the hard part, generally requiring millions of examples and millions of iterative adjustments. Once trained, however, the AI model provides answers nearly instantaneously.

 Figure 1 shows an AI model configured as a neural net. Input values go through layers of internal functions or "nodes", before being accumulated in a final answer. Although links are shown between just a few of the nodes, in many models each node is linked to all of the nodes in the previous layer, and provides results to all of the nodes in the following layer.

 Each node is a little calculator with adjustable variables that are carefully adjusted during training. For example, in "supervised" learning, the correct answer is already known ("ground truth"). The output is compared to the ground truth, and the internal variables are adjusted to obtain better agreement. After training, the model can be simplified by deleting unhelpful links and inputs, among other steps, resulting in a portable algorithm that solves problems fast, at negligible cost.



Fig. 1: Neural net AI model. Inputs are mathematically combined by layers of internal functions (nodes), which then feed the output. The ground truth is used during training to adjust variables.

 Figure 2 shows what's inside each node. Input links "X" are shown from the previous layer, and output links "Y" are shown going to the next layer. The internal function is actually quite simple: it calculates a weighted sum of the X inputs, and "squashes" the result between  $\pm 1$ . The weighting variables "W" and offsets "O" are adjustable. The squashing function is a trigonometric or logarithmic function. All the X inputs of a given node are different, but all the Y outputs of the node are the same. The final answer is the sum of all the node outputs from the last layer.

 Despite the primitive function in each node, trained AI models can provide surprisingly good answers. They are superb at finding subtle and



Fig. 2: A single node of the AI model. Node inputs from the previous layer are combined in a weighted sum using variable weights, then squashed and linked to the next layer.

complex correlations among thousands or millions of input values, which no human could possibly comprehend. The AI model, on the other hand, easily finds and huge, complex correlations. That's why AI often finds solutions that humans would never guess.

 It is apparent from the node structure that the AI model is not "intelligent" in any meaningful sense of the word. AI is just a massively-parallel basic calculator. Its only strength is to identify hidden correlations among a large number of input values, and to do so quickly (after being trained). Like any computer program, the AI model has no "will" or agenda of its own, despite the human-like appearance of some AI model outputs. But it's an illusion - those models were specifically trained to emulate humans. AI models only do what the operator trains them to do. AI can certainly do harm, like any powerful tool. But if AI does harm, blame the human operator, not the model.

#### Identifying and Localizing Message Faults

 Message faulting is an unsolved problem, and it is getting worse. Network crowding, pathloss at high frequencies, and the high numerologies and modulation orders desired for 6G all contribute to faulting. The current response to any message fault is to automatically request a retransmission of the message or its FEC bits (unless the message is already loaded with the FEC bits, a further burden). FEC bits sometimes work and sometimes not. Message faulting will be a serious time-waster in 6G, unless a better way can be found for correcting message faults.

 In a faulted message, there is still plenty of valid information remaining in the unfaulted symbols. Each faulted message element usually exhibits some kind of corruption signature, such as erratic modulation, unstable amplitude or phase, unexpected frequency shift or polarization angle, and other peculiarities. Shown below are waveform parameters that often accompany message faults. The receiver can identify the faulted message elements by detecting these signature parameters, and then correct the message using AI. Importantly, the entire fault correction can be implemented entirely within the receiver, without asking for a costly and time-consuming (and energy-consuming) retransmission.

 Correlating these parameters to determine the most likely faulted message elements is a complicated task, and correcting the message is even more so. But a trained AI model can easily analyze the disparate data, identify the likely faulted message elements according to waveform irregularities, discern the likely intent or meaning of the message based on prior unfaulted messages, and then provide the most likely corrected version - all in a tiny fraction of the time required for a retransmission.

 Figure 3 is an example of information that an AI model can use to diagnose a faulted message. The figure shows a modulation table with 16 states of either 16QAM (I versus Q branches) or 4x4 amplitude-phase modulation of the waveform signal. With either modulation scheme, or any other modulation scheme, the receiver can readily determine the modulation deviation of each message element, relative to the closest proper state. Most faulted message elements have a large modulation deviation, whereas correct message elements tend to be quite close. Two faulted message elements are shown in the figure as "o". The modulation deviation is either the absolute "radial" distance to the closest calibration state, or the cartesian coordinates of the two modulation parameters. The radial distance makes sense in QAM because the two branches are logically equivalent. The cartesian deviation makes sense in amplitude-phase modulation because the two axes represent different quantities, the waveform amplitude and phase.

 In amplitude-phase modulation, the transmitter modulates the waveform according to multiplexed amplitude and phase levels. The receiver still does the signal processing with orthogonal I and Q branches as usual, but then it calculates the waveform amplitude and phase using formulas. Demodulation is then done using the amplitude and phase values.

 Figure 4 shows another valuable waveform fault diagnostic. Using the digitized data of each symbol-time, the receiver determines the variations in the symbol amplitude, a clear sign of noise or interference. Noise-free signals are flat, other than the initial run-up. The amount of amplitude variation (within the subcarrier bandwidth) can be measured and correlated with faulting. The phase of a faulted signal is also likely variable, and even easier to detect.

 Figure 5 shows a distribution plot of the amplitude (or phase) variations, with and without noise, such as the noise shown in Figure 4. The distribution of amplitude variations is much wider when noise is present, as expected. The width and offset of the distribution depend sensitively on the type of noise, but in every case the width is increased relative to a noise-free signal.

 The AI model can take, as further inputs, the widths of the amplitude variation distribution and the phase variation distribution, and their offsets if any, for each message element in the message. The AI model can then identify the likely faulted message elements with highest waveform deviations and highest modulation deviations. The AI thus identifies each faulted message element in real-time, without a retransmission.



Fig. 3: The modulation deviation is the distance between the received signal and the closest calibration state. Faulted message elements have larger modulation deviations, on average.



Fig. 4: Amplitude variations during a symbol-time indicate noise or interference, and likely faulting.



Fig. 5: Distribution of amplitude variations during symbol-time with noise as in Fig. 4 (solid line) and without noise (dashed).

 The digitized waveform data of each message element can also reveal a small frequency deviation of each message element waveform, due to noise or interference. The frequency deviation is relative to the predetermined subcarrier frequency, but still within the subcarrier bandwidth. Message elements with a frequency offset are likely faulted.

 Many receivers can measure the polarization angle of the waveform using multiple antennas. Since the polarization angle is often affected by noise, any message element with a polarization angle different from the neighboring message elements is immediately suspect. The ratio of the two polarization signals is likely to be different in a faulted message element than an unfaulted one. The polarization ratio (or equivalently, the angle between the polarization components) of each message element is then another valuable diagnostic input to the AI model.

 Figure 6 shows an analysis of fault-correlation deviations. Some parameter of the signal waveform, such as amplitude or phase or polarization, is compared to the other message elements instead of the "nominal" or calibrated level. Any message element that deviates substantially from the other message elements is considered an outlier, and therefore suspicious.

 In the figure, the measured values of the waveform parameter are shown as dots for each message element. The distribution is shown as a peak. The width of the distribution shows the usual variation. The calibrated or "nominal" value of the parameter is also shown, as a line. The average often deviates slightly from the nominal value, but this is not a fault if they are all about the same. The "outlier", on the other hand, differs substantially from the average, indicating a likely fault. In the figure, the outlier is so close to the nominal value, that any regular test based on the nominal value would miss it. But in comparison to the average of the other message elements, the outlier is clearly different, and hence suspicious. To catch these outlier cases, the AI model can take, as further input, each message element's deviation from the average, for multiple parameters.



There are many other waveform parameters, indicative of faulting, that

the receiver can determine from the waveform data. It doesn't matter whether the data is still in the form of I and Q branch amplitudes, or has been converted to the waveform amplitude and phase values. In either case, the AI model can recognize the fault indicators of each message element, and calculates an overall "suspiciousness" metric that identifies the likely faulted message elements. Then the AI model, or another AI model, can correct the message, using the procedures described below.

#### Correcting the Faulted Message

 Identifying the likely faulted message elements is just a start. The AI model that identifies the faulted message elements, or a separate AI model, can be trained to correct faulted message elements based on various factors. For example, the AI model can use the information contained in the unfaulted message elements to determine the most likely correct version, or a number of candidate versions along with the likelihood of each one. All the candidate versions can be tested against the error-detection code associated withe the message (unless the error-detection code itself is faulted), among other tests described below. The AI model generally requires many inputs to discern the correct version of the message, but these are available in the digitized signal data. Figure 7 shows some of the parameters that the AI can use.

 The AI model can be trained to favor bit sequences or symbol sequences from previous unfaulted messages commonly received by the particular application at hand, and can strongly disfavor any sequences that are rarely or never seen in the prior messages. However, if one of the disfavored sequences is then received in an unfaulted message, that sequence can be added to the whitelist. In addition, sequences commonly seen in the faulted messages can be maintained in a blacklist, for further fault identification.

 The AI program can also recognize violations in rules, such as deviations from the accepted form of the message or its format. In addition, the AI can provide multiple candidate solutions, along with the likelihood of each version. For example, the AI can elevate candidates that alter only the likely-faulted message elements, and can downgrade candidates that alter non-suspicious message elements.

 For even greater value, the AI program can be trained to figure out the meaning or intent of the message, based on the unfaulted message elements, just a human expert would do. For example, the AI model can infer meaning according to prior unfaulted messages similar to the faulted one. The AI can also correlate the meaning or intent of the faulted message with current operating factors of the receiving entity, such as whether it receives an acknowledgement after a transmission, among many other, increasingly subtle, correlations.

 In a similar way, the AI model can select the candidate version that seems to "make sense" in the application, discarding or at least downgrading versions that seem inappropriate in the application. These judgements can again be based on prior unfaulted messages, as well as multi-parameter correlations that only AI can discern from the training examples.

 For further accuracy, the AI program can combine all the diagnostic results into an overall "suspiciousness" metric. This includes the number of message elements in the candidate that differ from the received message, and whether those altered message elements were likely faulted. It can also include factors such as whether each candidate obeys all form and format constraints, and whether the version corresponds closely to prior unfaulted messages, or includes rare or forbidden sequences, and whether the candidate version makes sense in the application, and whether it agrees with the error-detection code. The AI can then pick the candidate version with the lowest overall suspiciousness, as the corrected message.

 In a worst-case situation, the AI model can determine that there are too many faults to recover the message, or that the best candidate version still has high suspiciousness. In such cases, the AI model can recommend a retransmission of the whole message, or just a portion depending on the distribution of faults.

**7**

#### **AI INPUTS:**

#### WAVEFORM

 Amplitude fluctuations. Phase fluctuations. Received power level. Inter-symbol transitions. Polarization outliers. Frequency offset.

#### MESSAGE

 Modulation deviations. SNR each element. Error-detection code. Is EDC faulted?

#### EXPECTED

 Expected format, values. Past messages and faults. Past sequences. Rules and limits.

#### BACKGROUND NOISE

 Noise level during message. Interference level. Time structure of noise. Frequency structure. Signal in blank elements.

#### DEMOD REFERENCE

 Proximity to message. Both front and back? Mod quality of reference.

### RETRANSMISSION

 Also corrupted? Which elements differ? And how do they differ? Modulation deviations. SNR each element.



 If a retransmission is requested, and if it agrees with its error-detection code, and has the correct format, then the correct message is in hand, and that task is done. Then the AI, or another program, can diagnose the fault types in the corrupted message by comparing the modulations in the two copies. If, however, the second copy is also faulted, then the AI program, or another algorithm, can construct a merged version by selecting the message elements from each copy with the best signal quality and modulation deviation. The merged version usually has no faults, in which case the task is done. If the merged copy still has faulted message elements, the AI model can look for correlations in meaning between the first and second copy. Since the two copies have the same meaning but different faults, the AI can readily determine the corrected message, or at least a smaller set of candidate messages with greatly improved suspiciousness metrics.

 A major advantage of AI-based fault recovery is that the AI model can determine the most likely corrected version almost instantaneously, in a single pass through the neural net. Even for a complex inductive solution, based on meaning or intent for example, the correct version can generally be found in a tiny fraction of the time required for a retransmission. In most fault situations, the AI program can thus recover the correct message in a way that is completely transparent to the user. The AI has avoided the latency and dropped calls that users hate, and by recovering faulted messages, the AI has enabled the user's device to perform at the high level expected, even when the signal quality is poor. Ideally, the user is not aware that the message was initially faulted, and then rescued. On the other hand, a competitor's device, which lacks the powerful AI capability, would require two or three retransmissions to finally get the message right, at the expense of latency and energy consumption - assuming the link is not broken in the process. We have learned, from previous generations of wireless technology, that poor reception leads to customer dissatisfaction in a big way. The equipment with best message recovery always wins.

#### Selecting a Better Modulation Scheme

 Another important application of AI is to assist a base station in selecting a different modulation scheme when the fault rate gets too high. To select a better modulation scheme, the AI model can analyze fault types currently observed, and select a different modulation scheme with larger noise margins. The network can diagnose fault types by comparing the waveform amplitude and waveform phase of faulted and unfaulted messages. (Fault diagnosis is not feasible in QAM because noise scrambles the branches together, obscuring the cause.) Figure 8 shows the main fault types as: amplitude faults (Afault) which occurs when the amplitude is shifted by one amplitude level, phase faults (P-fault) when the phase is changed by one level, and non-adjacent faults (N-fault) when the amplitude or phase, or both, are off by more than one level. The base station or the user device can perform the fault type analysis, based on the uplink or downlink faulted messages that they have received.



Fig. 8: Fault types: Waveform amplitude, phase, and nonadjacent faults shown.

 After counting the rate of each fault type, the network can select a different modulation scheme to mitigate the fault types observed. The selection of a particular modulation scheme in a busy network environment is a complex process due to the huge number of possible modulation choices (see below) and the many competing interests such as high throughput, fault minimization, low latency, retransmission avoidance. A trained AI model can perform this task instantaneously and prevent further faulting.

 Many modulation schemes are available to networks, each with different margins, costs, and capabilities. The most beneficial modulation schemes at high frequencies are based on amplitudephase modulation, because amplitude-phase modulation provides larger noise phase margins than OAM. For example, figure 9 shows an "asymmetric" modulation scheme, in which the number of amplitude levels is different from the number of phase levels. Here Namp=8 and Nphase=2, thereby providing 16 states. This is the same number as 16QAM, but now every state has a full 180 degree phase margin (measured between centers). Unlike QAM, this modulation scheme would eliminate phase faults at high frequencies. A brief demodulation reference may be placed at the start and/or end of each message, if necessary to discriminate the eight amplitude levels. When phase faulting becomes problematic with 16QAM, as it will, the network can switch to the depicted modulation scheme and the phase faults will vanish. At low frequencies, on the other hand, when amplitude faulting is more prevalent than phase faulting, the network can use a modulation scheme that has more phase levels and fewer amplitude levels. Asymmetric demodulation is not feasible in QAM because the I and Q branches are logically equivalent.

 Figure 10 shows another beneficial modulation scheme. Here the number of amplitude levels is 6 and the number of phase levels is 3, neither of which is a power of 2. The scheme provides18 states, each with 120 degrees of phase margin. The scheme virtually eliminates phase faults, while providing higher throughput than 16QAM, due to the two extra states. Also shown are "acceptance regions" around each modulation state, such that any modulated element falling within one of the acceptance regions is automatically demodulated according to the associated modulation state. Also shown are "exclusion zones" such that any element with modulation in one of the exclusion zones is automatically flagged as faulted. The acceptance regions can be tailored to the current noise environment. In this case, tighter limits are imposed on on amplitude than on phase, as seen by the oval shape.

 Figure 11 shows another amplitude-phase modulation scheme with 16 states, but now the amplitude levels are not spaced uniformly. A larger spacing is provided at the low-amplitude scale, and smaller spacing at the high-amplitude end. This is to compensate for the relatively low SNR at low amplitudes. The phase separation is 90 degrees for every state, unlike QAM.

 The network must consider many factors before deciding whether and how to switch modulation schemes. The network must also consider the modulation order, the numerology, and other conventions such as repetitions, and other variables that further broaden the range of available modulation choices. The network must also consider the QoS and QoE of each user device,



Fig. 9: Asymmetric modulation. Namp=8, Nphase=2, Nstates=16



Fig. 10: Non-power-of-2 modulation with acceptance and exclusion zones.



Fig. 11: Variable amplitude spacing to optimize reliability.

which may be parsed further to prioritize low latency versus high throughput, and overall reliability versus a tolerance for some faulting (which the receiver may be able to correct, as mentioned). Further considerations include whether the faulting occurs in certain frequency bands, or at certain time intervals, in which case the network can switch to a different frequency and/or transmission schedule to avoid the observed time/frequency noise bursts. Some users with strict latency requirements cannot use retransmissions at all, because they arrive too late; only the original messages arrive on time. Some users are more flexible regarding latency but depend on the message finally being corrected. In crowded networks, priority can be placed on minimizing unnecessary transmissions and minimizing transmission power to avoid background generation, but not so low that signal quality suffers since that would result in more faulting and more transmissions and more crowding. Battery-constrained devices may prefer avoiding retransmissions in their uplink messages, and in fact may wish to apply even higher uplink transmission power just to avoid frequent retransmission requests by the base station. Different modulation schemes can be provided for uplink and downlink, and different combinations for each user device. Almost every one of these parameters is a compromise between multiple competing interests, all of which change dynamically as the background fluctuates and the current demand changes. Human operators cannot possibly assess each episodic problem, decide on a mitigation such as a modulation change, and implement it before the next conflict emerges. Therefore, in every fast-cadence 6G base station, AI is needed for network management.

 Complex nonlinear problems such as this, with multiple competing interests, each with different priorities, are ideal for AI. In fact, a well-trained AI model may come close to optimizing the overall performance, as viewed by the users and also by the network operators. For example, the network can provide each user's priorities, its current fault spectrum, and its computational capabilities, as inputs to the AI model. The model then indicates which modulation scheme is best suited to each user device, for uplink and for downlink. The AI program can also include, in the calculation, the non-negligible cost of switching modulation schemes, such as the added communication costs required to inform the affected user devices.

 After the AI has demonstrated high competence at selecting modulation schemes to optimize each user's communication experience, the network may decide to turn over responsibility for modulation control directly to the AI model. The AI then autonomously handles the entire process, including selection of the modulation scheme, transmitting the necessary change alerts, and monitoring the results. Thus the network parameters, such as modulation for each user device, can be controlled automatically, in real-time, without human intervention. This improvement would result in smoother network operations, lower network costs, and improved customer experience overall.

#### Adjusting Beam Parameters

 Another important application of AI is adjusting downlink beam parameters, such as direction, width, frequency, power, and polarization, for optimal reception by each user device. The transmission beam properties are influenced by numerous competing interests, such as high reception reliability, minimal energy consumption, minimal background generation, and low latency. These priorities are generally different for each user device, and different still for uplink. The best compromise usually depends on many environmental factors such as the noise and interference experienced by each user device, including the frequency and time distribution of the interfering signals, the distance of the receiver from the transmitter, and the density of user devices in the beam direction (regarding background sensitivity). The best compromise also depends on the priorities of the receiving entity, including the QoS and QoE of the entity, but with greater granularity in terms of latency, reliability, signal quality, message size, computational demands, whether the transmitting or receiving entity goes off-line

periodically to save energy, whether the receiving entity is an emergency user or other escalated-priority user, and in some cases whether the user has purchased enhanced priority.

 AI is perfect for situations like this. There are too many factors, and too many competing relationships among them, for any human to fathom, and conditions change far too quickly for any human to react in real-time. A well-trained AI model, on the other hand, can slice through the complexity and indicate the best transmission beam parameters, in mere microseconds. For example, the base station can include an AI model that takes as input the current and anticipated downlink message load, the QoS and other preferences of the user recipients of those messages, and the current background/interference conditions reported by various user devices. The AI can thereby provide beam parameter settings customized for each user's needs, and optimized overall.

 Ultimately, the AI model may be able to select and implement these beam control changes autonomously. Preferably such autonomous operation may be permitted only after the AI program has demonstrated good performance and stable operation for extended intervals, without human intervention. Using AI to "close the loop" in this way enables near-instantaneous reaction to changing conditions, on a time scale that would be impossible if humans are in the loop, due to the time required for the human brain to comprehend the changing conditions and make some kind of adaptation, but also because the inter-related priorities are just too complicated.

 The reception beam parameters of the base station can also be controlled by AI, using a different set of inputs but basically the same model. Since the base station must receive messages from multiple sources simultaneously in each OFDM symbol, the AI can optimize the phase and gain properties of each antenna in real-time according to the angles of the various transmitting users at each time.

#### **Conclusions**

 AI is transforming technology, and especially wireless technology. In this whitepaper, a few prominent applications of AI are outlined, leading to improved beam control for optimal reception, improved modulation schemes to mitigate faults, autonomously correcting message errors by the receiver without a retransmission, and determining the meaning or intent of the message despite corruption.

 Wireless developers and producers should recognize AI as a business opportunity to gain competitive advantage, as opposed to setting standards. For example, a company installing AI-based fault correction in their receivers can offer their customers enhanced message reliability and fewer dropped calls, without the latency and energy costs of a retransmission. A company that produces base station electronics with AI-based modulation selection and autonomous beam optimization can provide substantially improved communications to the users at reduced energy costs. In these cases and many others, AI directly provides improvements in performance, which leads to improved customer satisfaction. Since customer satisfaction drives market share, companies planning the transition to AIcentric operations have a unique opportunity to lead. AI is the key that opens all these doors.

### Glossary

"Base station", as used herein, includes all network assets communicating with users, including access points, access relay stations, roadside monitors, satellite relays, and the like. The term also includes the core network, backhaul, and other internal systems of the network assets, unless otherwise called out.

"User device", as used herein, refers to the radio portion of user equipment, specifically the transmitter, receiver, antenna, signal processing electronics, and demodulation processor. The term also includes AI models for fault mitigation and message interpretation and the like, when present.

3GPP (Third Generation Partnership Program) is the primary organization for wireless technical specifications, and with seven "Partner" organizations, promulgates universal wireless standards.

OFDM (Orthogonal Frequency-Division Multiplexing) means transmitting message data in multiple frequencies (subcarriers) at the same time. The receiver then measures the subcarrier signals to separate and demodulate the message elements.

IoT (Internet of Things) devices are low-cost, reduced-capability wireless sensors and actuators.

SNR (Signal-to-Noise Ratio), as used herein, includes interference, stochastic noise, clock drift, and all other effects causing message faults, unless specifically indicated.

FR1 and FR2 are frequency ranges. FR1 is 7.125 GHz and below (and up to 8.4 GHz in 6G). FR2 is 24.25 GHz and up. FR2 is often called mmWave, although a wavelength of 1 mm actually corresponds to a frequency of 300 GHz.

BPSK (binary phase-shift keying) is phase modulation at constant amplitude with 2 states separated by 180 degrees, carrying 1 bit per symbol.

QPSK (quadrature phase-shift keying) is phase modulation at constant amplitude with 4 states separated by 90 degrees, carrying 2 bits per symbol

QAM (Quadrature Amplitude Modulation) is a modulation scheme in which the message data is encoded in the amplitudes of two orthogonal signal components, termed I and Q branches.

A resource grid is an array of resource elements, arranged by symbol-times in time and subcarriers in frequency.

A message element is a single modulated resource element of a wireless message.

A "symbol-time" is the time duration of a single message element.

A message is "time-spanning" if the message elements are sequential in time on the same subcarrier, and "frequency-spanning" if the message elements are sequential in frequency at the same symbol-time.

PDSCH and PDCCH represent the downlink shared and control channels by which the base station communicates with each user device.

### References

[1] The following artificial intelligence patents can be found at: www.UltraLogic6G.com.

