Machine Learning for Ambient Backscatter Channel Estimation and Signal Detection: Opportunities and Challenges

Diancheng Cheng, Fan Wu(✉), Cong Zhang, Jinjin Sun, and Yuan'an Liu

Beijing University of Posts and Telecommunications, Beijing 100876, China
wufanwww@bupt.edu.cn

Abstract. As a promising low-power connection paradigm in the ubiquitous Internet of Things (IoT), ambient backscatter communication (AmBC) collects energy from ambient radio frequency (RF) signals while using them as carrier signals, which brings ultra-low power consumption and deployment cost. However, it has not been widely applied in practice because of its difficulties in weak signal detection. To overcome these difficulties, machine learning (ML)-based methods have been highlighted recently. ML methods can achieve accurate signal processing under a low receive signal-to-interference-plus-noise ratio (SINR) in unpredictable interference communication scenarios, benefiting from their outstanding inference and classification tools. In this survey, a brief review of AmBC is first introduced and the four-fold signal-receiving challenges of AmBC are discussed. After that, two key signal processing technologies, i.e., AmBC channel estimation and AmBC signal detection, are emphasized. The representative ML-based methods of AmBC channel estimation and AmBC signal detection are summarized, following their advantages and disadvantages. Finally, some valuable research directions on this topic are introduced to guide future research.

Keywords: Ambient Backscatter Communication · Machine Learning · Channel Estimation · Signal Detection

1 Introduction

Backscatter communication, e.g., Radio Frequency IDentification (RFID), introduces an energy-saving connection paradigm to the Internet of Things (IoT), which uses load modulation to re-modulate dedicated or ambient carrier signals in a passive manner instead of generating radio waves actively. Because of no need for bulky and costly components (e.g., mixer, oscillator, ADC, etc.) on backscatter devices (BDs), such a scheme is particularly suitable for ultra-low-power and low-cost IoT scenarios. There exist three architectures of backscatter communication systems, including monostatic backscatter communication (MoBC), bistatic backscatter communication (BiBC), and AmBC [1]. Among

X. Li et al. (Eds.): SmartGift 2024, LNICST 600, pp. 3–19, 2025.
https://doi.org/10.1007/978-3-031-78806-2_1

them, AmBC is regarded as one of the most promising backscatter architectures since it has the potential to be deployed in an existing wireless network arbitrarily and backscatter data in the same frequency band with ambient RF sources, which will significantly reduce deployment costs while increasing spectrum efficiency. However, AmBC has not been widely applied in practice yet because of its crucial technical challenges in signal receiving. As the price of using free RF resources in the air, AmBC needs to work with uncontrollable RF sources and unpredictable carriers. Besides, the backscattered signal from the BDs is generally much weaker than the direct-link signal from the RF sources because the backscattered signal suffers from double path loss. Thus, it is more difficult to recover useful information than traditional wireless communication systems.

To solve signal-receiving challenges, numerous research on AmBC signal processing have been proposed, which mainly focus on two key technologies: AmBC channel estimation and AmBC signal detection. These works were supposed to pave economic and efficient ways for AmBC receivers to estimate channel parameters and distinguish symbols with little prior knowledge (e.g., carriers and channel state knowledge) and little cooperation between ambient RF sources and passive BDs. Many simple and classic methods were extended to AmBC scenarios in the early stages. For example, blind channel estimation [2] and energy detector [3] provided the most intuitive solutions to AmBC channel estimation and signal detection but with limited performance. Differential encoding [4] can release the need for channel estimation, but increase the power consumption of BDs. The classic maximum-likelihood estimation [5] was suitable for AmBC signal detection but needed a complex DLI cancellation design before detection.

Recently, machine learning (ML)-based methods have introduced new solutions to many intractable wireless communication problems without needing exact mathematical models or explicit programming, which can achieve outstanding performance [6]. As for AmBC channel estimation and signal detection, which is difficult to acquire necessary prior knowledge for modeling and is significantly differences from deterministic signal processing, ML-based methods have the potential to overcome the challenge of lacking knowledge of ambient RF source and channel state information (CSI) and reduce costs for canceling direct-link interference (DLI) from the ambient RF sources.

In this paper, we first provide a brief review of AmBC from the perspectives of history, architecture and working paradigms, which explains the advances of AmBC compared to other backscatter communications. Meanwhile, we clarify the crucial backscatter signal receiving problem of AmBC from four aspects and then highlight two key signal processing technologies: AmBC signal estimation and AmBC signal detection. After that, we emphasize ML-based methods for AmBC channel estimation and AmBC signal detection because they often outperform traditional methods benefiting from powerful classification, clustering, and neural network algorithms. We discuss the unique applicability of ML methods for AmBC signal receiving, and introduce ML methods for the two key technologies, respectively. In these sections, we summarize the goals, technique challenges, classical solutions, and current novel ML-based solutions according to different ML algorithms. Specifically, we classify these ML-based solutions into two categories according to whether they conduct feature engineering in advance or not. We discuss their methodology, advantages, and disadvantages. Finally, we propose some future research directions for ML methods on AmBC systems standing in the perspective of technique processes.

2 Brief Overview and Signal Receiving Issues

2.1 History, Architectures and Working Paradigms of AmBC

AmBC was first proposed in 2013 [7], which has been regarded as a promising technology to achieve ultra-low-power and low-cost communication in a passive manner. It can assist existing wireless communication networks (e.g., cognitive radio networks, wireless powered communications networks, etc.) to promote spectrum efficiency as a secondary system or be applied on machine-type communications (e.g., massive IoT networks) as a low-power connective scheme [1]. After the birth of AmBC, various ambient RF signals were explored as potential sources and carriers to realize ambient backscattering, starting with digital television (DTV) signals [7] and then extending to FM [8], Wi-Fi [9], Bluetooth [10], and even long range (LoRa) signals [11] (shown in Table 1). These explorations shaped the prototypes of AmBC systems and proposed many basic enabling technologies, e.g., load modulation, frequency shifting, etc. However, these prototypes can only work in short range and with low data rates limited by their architecture and working paradigms.

Table 1. Prototypes and Achieved Perfomances of AMBC Systems

Systems	Sources	Data Rate	Range	Pow.Consumption
Ambient Backscatter [7]	DTV	1 kbps	0.46 m–0.76 m	0.25μW
FM Backscatter [8]	FM	3.2 kbps	1.5 m–18.3 m	11.07μW
WiFi Backscatter [9]	WiFi	1 kbps	2.1 m	0.65μW
BLE Backscatter [10]	BLE	16.6 kbps	25 m–65 m	37μW
PLoRa [11]	LoRa	1.58 bps–2.23 bps	1.1 km	220μW

AmBC differs from other backscatter communications in architecture and works in three typical paradigms. Among the three architectures of backscatter communications, as shown in Fig. 1 (A), both MoBC and BiBC have dedicated carrier emitters, so they can control the whole process of backscattering autonomously. Especially when acting on signal receiving, their receivers have full prior knowledge of the carrier signal. Differently, AmBC captures ambient RF signals as carriers, which are uncontrollable and unpredictable. Therefore, signal receiving in AmBC is not a determinate signal processing problem anymore. In

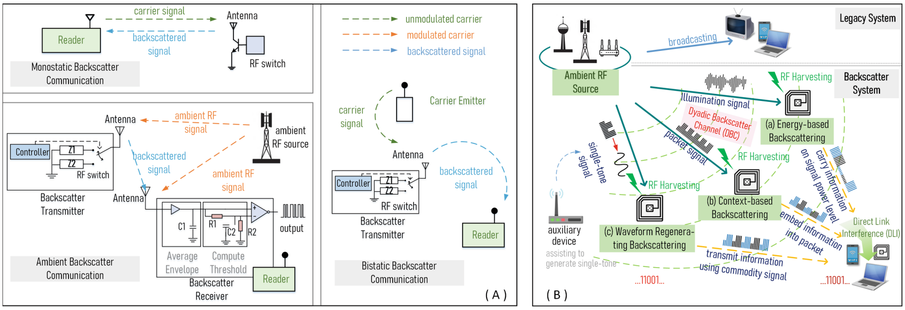

Fig. 1. A: Three architectures of backscatter communication. In MoBC, the carrier emitter and receiver are co-located at the reader. In BiBC, the carrier emitter is physically separated from the reader. In AmBC, the carrier signal is from an ambient RF source and there is no dedicated carrier emitter. **B:** Three working paradigms of AmBC. In energy-based backscattering, the ambient RF signal is regarded as an illumination signal and the BDs carry information on the signal power level. In context-based backscattering, generally, the packet signal is used as a carrier signal to embed information into the packet. In waveform regenerating backscattering, the single-tone signal that is purified from the ambient RF signal or generated by the auxiliary device is exploited to re-generate a waveform that is compatible with commodity devices.

reality, AmBC has three typical working paradigms according to current research work, including energy-based backscattering, context-based backscattering, and waveform regeneration backscattering, as shown in Fig. 1 (B). They carry useful information onto ambient RF signals in different piggyback manners. No matter which kind of working paradigm, however, signal receiving in AmBC faces serious DLI and low receive SINR, which are not so critical in traditional wireless communications and other backscatter communications.

2.2 Signal-Receiving Challenges in AmBC

The architecture and working paradigms make AmBC work at a low data rate (typically from several bps to several kbps) and a short communication range (typically from dozens of centimeters to several meters). Compared with mature backscatter communication, e.g. RFID, there is still a performance gap. We discuss the challenges of AmBC signal receiving from a system perspective and summarize them in the following four aspects.

1. **The Ambient RF Source:** The ambient RF source is uncontrollable, and the ambient carrier is unpredictable. The AmBC systems do not deploy their dedicated carrier emitters but capture and re-modulate RF signals from the surrounding open air. This is the reason why AmBC can cost much less on deployment than other backscatter systems and has the potential to achieve "plug and play" in existing wireless networks. However, these ambient RF sources are intended for legacy systems rather than AmBC systems. They

will not cooperate with AmBC systems and even cannot perceive their existence. Both BDs and ambient receivers have no sufficient knowledge about ambient RF sources used for carrier recovery. Besides, the ambient carrier is unpredictable since it is easily affected by the open environment.

2. **The BDs:** The BDs are limited in function due to strict power constraints and scarce hardware resources. Most BDs are powered by energy harvested from RF signals, which is typically at a μW level. They generally just act with load modulation and limited digital control. Since BDs work in a passive manner, they are not equipped with active RF components, e.g., amplifier, ADC, etc. Therefore, they cannot amplify incident signals, as well as signal processing and control. After energy harvesting and re-modulating at BDs, the modulated backscatter signals become weaker and more difficult to distinguish.
3. **The Backscatter Receivers:** The receivers are complex to design and confronted with an extremely low receiving SINR. With uncooperative ambient RF sources and simply functional BDs, the ambient receivers have to take on almost all the responsibility of communication control, which requires complicated hardware and protocol design. What interests more in this paper is that the ambient receivers are supposed to recover useful information under a very low receiving SINR, because the signal from direct link is generally several orders of magnitude higher than the backscatter signal in an AmBC system [4]. The receivers should adopt efficient DLI cancellation and weak signal-receiving technologies.
4. **The Channel Model:** The backscatter channel in AmBC has deeper fading. The modulation backscatter channel, also called the Dyadic backscatter channel (DBC), has different statistical characteristics from traditional one-way channels, and it has deeper fading [12]. This channel model mainly contains two paths of signals: the ambient RF signal from the direct link and the useful signal from the backscatter link. Since the useful signal travels from an ambient RF source to a BD and then is backscattered to a receiver from the BD, it suffers from double attenuation, which makes backscatter signals fade deeper.

2.3 Key Signal Processing Technologies of AmBC Signal Receiving

To overcome all the aforementioned four-fold challenges in AmBC signal receiving is a systematic problem. It involves a lot of hardware and software technologies, e.g., on-tag circuit designs, DLI cancellation, channel estimation, signal detection, etc. Among them, AmBC channel estimation and signal detection have become hot points these years because they are quite common in AmBC signal receiving and play a key role in signal-receiving challenges.

1. **AmBC Channel Estimation:** Channel estimation is an essential part of signal receiving, which can provide the necessary parameters for signal detection. Further, it can provide key information about instantaneous or statistical channels for transceivers to make communication more efficient and safe. According to the DBC model, there are three individual channel parameters,

which can be seen in Fig. 1 (A), that need to be estimated to obtain full channel state information (CSI). Once the perfect CSI is known, different useful information about BDs can be calculated according to received signal models. More commonly, estimating the partial channel parameters of a direct link and cascaded backscatter link is sufficient to distinguish different backscatter symbols. However, channel estimation is not easy in AmBC, since neither the ambient RF source nor BDs can provide training pilots. Blind channel estimation [2] was proposed naturally, but its performance was poor. Non-blind channel estimations are supposed to provide better estimation performance [13–15], but the cost may not be acceptable in such a power-constrained backscatter system.

2. **AmBC Signal Detection:** Signal detection in AmBC mainly refers to symbol information detection without perfect CSI, as well as knowledge of the ambient RF source. Since ambient channel estimation has been a costly task so far, numerous current studies pave economic ways to directly recover symbol information without estimating channel parameters [16–20]. Despite lacking knowledge of ambient RF source and channel parameters, the backscatter symbols are distinguishable for the receiver because the load modulation at BDs will introduce different link paths. When BDs reflect incident RF signals, there are both a direct link and a backscatter link; but when BDs absorb incident RF signals, there is only the a direct link. The ambient receivers can distinguish different symbol information by exploring power levels or other unintuitive features of the received signals. When it comes to distinguishing different categories with incomplete signal models, unknown prior knowledge, and implicit features, machine learning methods are highlighted.

3 Machine Learning Methods for AmBC Channel Estimation and Signal Detection

3.1 The Trends of Machine Learning Methods

Different from traditional wireless communications that are well-designed to achieve precise system cooperation with perfect prior knowledge, AmBC has difficulties extracting and recovering interested symbol information from a low SINR received signal without CSI due to the aforementioned four-fold challenges. Statistical signal processing technology is favored in AmBC because of lacking key parameters of the received signal. Recently, machine learning methods, which have rich experience in feature extraction and classification, have become a hot research point in AmBC channel estimation and signal detection. Here follows its unique applicability in AmBC.

1. **High fitness for stochastic signal processing tasks in AmBC:** The received signal in AmBC is of great stochasticity for ambient receivers since it comes from an unknown source and experiences unknown re-modulation. The statistical learning methods, e.g., probability analysis and parameter estimation, are often preferred to deal with stochastic signals. As a classical statistical learning method, machine learning can effectively assist AmBC receivers

to obtain channel parameters from the received signal without empirical knowledge [13,15] and distinguish different symbol information [16–20].

2. **Performance improvement on AmBC channel estimation and signal detection:** The existing methods of AmBC channel estimation and signal detection mainly explore the backscattered signals from the perspective of communication and often focus on easily observed intuitive physical quantities (e.g., energy level), which will limit the performance to system design experience. The ML-based method is a data-driven method. It is skilled in dealing with rich and complex signal characteristics in data, not limited to intuitive physical quantities (e.g., signal constellation [19]), which can improve the accuracy of AmBC channel estimation and signal detection. In addition, the deep learning method can often mine more effective features beyond design experience through a specific neural network (e.g., the spatial and temporal correlation of the received pilot signal [15] and the eigenvalue of the sampling correlation matrix [20]), which will significantly improve the performance.
3. **Adaptable to volatile and massive communication scenarios:** AmBC captures carriers from open air, which is vulnerable to the legacy wireless network and surrounding environment. When they change, the methods based on mathematical analysis and specific models cannot be inferred and predicted, which brings serious performance degradation. In addition, when an AmBC system comes to massive IoT scenarios, the channel and signal model becomes quite complicated, which makes it hard to propose an accurate model and obtain a mathematical solution. Fortunately, data-driven ML-based methods have inference and prediction ability. They can adjust the trained model flexibly and timely to improve the performance of AmBC channel estimation and signal detection in these scenarios.

3.2 Machine Learning Methods for AmBC Channel Estimation

In AmBC, many researchers assume perfect CSI is known before signal detection. However, estimating channel parameters in such a low-power backscatter communication system is challenging and unaffordable. This section will discuss specific goals and technical challenges of AmBC channel estimation, and introduce ML-assisted methods, which learn implicit channel parameters from noisy training pilots.

Goals: A typical three-node AmBC system is composed of an ambient RF source, a single-antenna BD, and a single/multi-antenna receiver, as shown in Fig. 2 (A). It usually fits the following channel models:

– There are mainly two links: A direct link from the ambient RF source to the receiver (S-R) and a double attenuated cascade link relayed by the BD (S-BD-R).
– When transmitting symbol '0', the BD absorbs all ambient signals without backscattering, so there is only a direct link. While transmitting '1', the BD backscatters ambient signals introducing the cascade link.

– Since BDs are very close to ambient RF sources for collecting more energy, the channel between them can be usually regarded as a constant.

The goal of AmBC channel estimation is to obtain the individual or combined channel parameters in the above channel model when transmitting different symbols. It is mainly used to recover symbol information since the channel changes with load modulation at BDs.

Technical Challenges: Channel estimation in AmBC is much more difficult than that in traditional wireless communication. The commonly used channel estimation methods in traditional wireless communication include blind estimation, semi-blind estimation, and non-blind estimation. Non-blind estimation algorithms based on pilot training are most commonly used, e.g., least square (LS), linear minimum mean square error (LMMSE), and minimum mean square error (MMSE). They occupy a certain spectrum of resources but have high accuracy. However, they are not applicable in AmBC. The reasons are as follows:

– Since the ambient RF source is not cooperative, it cannot provide necessary training pilots, and BDs cannot transmit pilots actively either due to energy and hardware constraints.
– The channel parameters are not consistent when the BDs transmit different symbols since the channel changes when BDs act load modulation, which increases the complexity of estimation.

Although blind channel estimation methods can be used to avoid the need for pilots, their performance is not appealing. The authors in [13] combined the expectation maximization (EM) algorithm to blind channel estimation in a typical AmBC system. However, it can only obtain the amplitude of the channel parameters, which is far from meeting the requirement of accurately describing AmBC channels. Therefore, non-blind and semi-blind channel estimation is still the research hotspot for AmBC channel estimation [21]. AmBC channel estimation needs to solve the problems of obtaining training pilots and estimating inconsistent channel parameters when BDs transmit different symbols.

Protocol for Passive Pilots Acquirement: To obtain training pilots, a classic communication protocol is widely used in the AmBC system, as shown in Fig. 2(B). In the transmission frame, it utilizes the amount of symbols with known contents as training pilots before transmitting useful data, which replaces the need for necessary pilots from the ambient RF source. Considering transmission synchronization in this protocol, only simple interaction between the receiver and the BD is needed in practice. At the beginning of each frame, the receiver will send a high-level pulse to inform BD of the transmission start [20].

Besides, to solve the problem that channel parameters vary with BD symbols, it is necessary to provide an individual training phase for every symbol in the above transmission frame. In current studies, which prefer simple binary

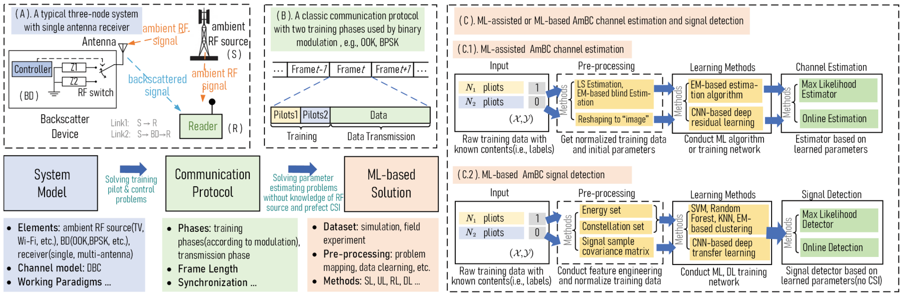

Fig. 2. The structure and process of machine learning method for AmBC channel detection and AmBC signal detection.

modulation (e.g., OOK and BPSK), there are two training phases for estimating inconsistent channel parameters before transmitting useful data. It is worth noting that this popular protocol has a trade-off between training accuracy and data rate since it has to provide more training phases while the BD adopts high-order modulation. However, it is still the mainstream solution because of its simplicity, and we highlight this protocol because it has been widely used in ML-based methods in AmBC channel estimation and signal detection.

Current ML-Assisted AmBC Channel Estimation: After solving pilot problems, non-blind and semi-blind channel estimation methods can be applied in AmBC. The traditional methods make it difficult to achieve accurate and efficient performance in AmBC. For example, LS usually regards channel parameters as deterministic but unknown constants, but the AmBC channel is a random variable affected by the environment, so the estimation is not accurate. Although LMMSE and MMSE regarded the channel parameters as random variables, they need quite a lot of computation to determine a model. In addition, when the channel changes, it is expensive to recalculate a new model. To promote estimation performance and provide a model with inference ability, ML-assisted methods are favored in AmBC channel estimation, which is summarized in Table 2. These methods are introduced to classify channel vectors or matrices into different categories according to different symbols without sufficient prior knowledge.

To apply machine learning methods in AmBC channel estimation, the general process is shown in Fig. 2(C.1). The input raw data is sampled from training phases of the classic communication protocol (generally with unknown content as labels). Before the training task, the raw training data must be pre-processed to obtain initial parameters or get a normalized input format for training the model. Then, appropriate machine learning algorithms are chosen for model training. Finally, the estimated channel parameters are applied to corresponding AmBC systems, and the well-trained model could be deployed in corresponding online transmission.

Table 2. ML-Based Methods for AMBC Channel Estimation

Category	Methods	Advantages	Disadvantages
ML-assisted channel estimation	UL (EM, semi-blind channel estimator) [13]	1. Obtain combined channel parameter 2. More efficient than blind channel estimation	1. Need few pilots from RF source 2. Need rough estimation
DL-based channel estimation	DL (CNN, Residual learning,denoiser) [15]	1. Obtain channel coefficient vectors 2. Hign performance close to optimal MMSE channel estimation	1. Need larger dataset 2. Higher computational complexity

In [13], the authors proposed a semi-blind ML-assisted methods based on the EM algorithm. They used LS estimation to pre-process pilots and obtained rough channel parameters as initial parameters of EM estimation. It achieved more accurate channel state information than blind channel estimation. In the deep learning method, the channel estimation problem can be transformed into an image denoising problem, which has achieved many mature solutions [15,22]. In the pre-processing phases, the training pilots are reshaped into two-dimensional data, and a noisy channel "picture" is obtained. Then the deep learning image denoising method is used to obtain a well-denoised picture. Finally, the picture is recovered to a one-dimensional vector to obtain the AmBC channel parameter. In [15], the authors proposed a deep residual learning method based on the convolutional neural network (CNN) to estimate the AmBC channel. The proposed architecture denoised the channel picture through several denoising units, each of which contained an L-layer network. The method can achieve an accuracy comparable to the optimal MMSE channel estimation under perfect assumption. That is because the temporal and spatial characteristics of the pilots are utilized in the process of denoising the channel picture.

3.3 Machine Learning Methods for AmBC Signal Detection

In most traditional wireless communication, signal detection is a series of signal processes of recovering useful information under demand received SINR after perfect channel estimation. However, the received SINR in AmBC is quite low, and perfect CSI is difficult to obtain. ML-based AmBC signal detection introduces effective and economical manners.

Goal: The goal of AmBC signal detection is to recover different modulated symbols from the received signal under extremely low SINR without perfect CSI. It is worth noting that binary impedance modulation is widely used in current research because of its simplicity. Therefore, the goal of AmBC signal detection is usually to detect two different symbols in the superimposed backscatter signal.

Technical Challenges: AmBC signal detection is significantly different from traditional wireless communication:

– AmBC receivers face weak signal detection problems under ultra-low SINR because the received backscatter signal is generally several orders of magnitude lower than the ambient RF signal from a direct link. DLI is desperately needed.
– The BDs are low-power passive devices, which do not have sufficient energy and hardware resources to support complex signal processing and control. Therefore, they cannot amplify incident signals and act in complex modulating and encoding, which makes the useful backscatter signal hard to distinguish.
– The CSI is absent. AmBC channel estimation is a challenging task. The receiver cannot know the knowledge of ambient signals and channel state, which makes AmBC signal detection more difficult.

Classical Solutions: In the early state after AmBC was proposed, many simple and intuitive non-coherent detectors were proposed, including energy detection (ED) [7] and maximum likelihood detection (MLD) [5].

The ED is based on received signal power levels between different symbols. It is widely used in early prototypes, and it can be implemented only using analog circuits. However, it is merely suitable for a very short distance. In [7], the authors designed a classic analog ED. It averaged and smoothed the envelope of received backscatter signal first, and then detected symbol information by a carefully designed threshold comparing circuit. However, in the superimposed received signal, the useful backscatter signal is often drowned in the strong direct link ambient RF signal, which significantly reduces its detection performance.

The MLD is based on received signal statistical distributions between different symbols, which regards AmBC signal detection as a hypothesis testing problem [5]. Such algorithms often assume that the ambient RF signal follows circularly symmetric complex Gaussian distribution (CSCG) distribution and the superimposed received signal will follow the same statistical distribution model (e.g., Gaussian Mixture Model) with different parameters when BDs backscatter different symbols. By analyzing the received signal, the likelihood or log-likelihood function can be designed to estimate different distribution parameters in the hypothesis space. However, the real statistical distribution model in AmBC is more complex than the assumption above since the surrounding environment is changeful, and the differences between assumed models are usually not significant because the ambient RF source performs a dominant contribution instead of a useful backscatter signal.

Current ML-Based AmBC Signal Detection: Compared to ED and MLD, ML-based AmBC signal detection can conduct more distinguishable features between different symbols, and has powerful classification and clustering tools to put received signals into correct symbol categories. From the perspective of constructing features, ML-based methods for AmBC signal detection can be divided into ML methods based on feature engineering and deep learning (DL) methods based on the artificial neural network (ANN), as summarized in Table 3

Table 3. ML-Based Methods for AMBC Signal Detection

Category	Methods		Advantages	Disadvantages
ML Methods Based on Feature Engineering	Energy based	SL (SVM, random forest) [16]	1. Jointly consider channel estimation and signal detection 2. Directly recover symbol without CSI	1. Difficult to collect labeled dataset 2. Need to cancel DLI
		SL(KNN) [17]		
		UL(EM) [18]		
	Pattern based	UL(EM) [19]	1. Little affected by DLI 2. Fast convergence rate	1. Need to analyze input feature 2. Need remarkable pattern feature
DL Methods Based on ANN	DL (CNN, DNN, transfer learning) [20]		1. More distinguished feature(e.g., temporal and spacial features) 2. Higher detection performance	1. Need larger dataset 2. Hign computational complexity

ML Methods Based on Feature Engineering: These methods need to construct feature engineering according to specific problems before model training. The features describe the individual characteristics of the input data, and feature engineering is generally conducted according to background knowledge and previous experience. In AmBC, the popular features include intuitive physical quantity [16–18] and communication patterns [19] of the signal. At present, a large number of ML algorithms have been applied to AmBC signal detection based on feature engineering. Technically including, supervised learning (SL) methods, e.g., support vector machine (SVM) [16], random forests [16], K-Nearest Neighbors (KNN) [17], and unsupervised learning (UL) methods (e.g., EM algorithm [19]).

The SVM and random forest methods are explored in [16], where a typical AmBC system with BPSK modulation at BD was adopted, and it assumed system deployment and channel states remained unchanged. This work utilized energy sets as input features, which were conducted from the average of N signal samples after data enhancement. Since the contents of signal samples were known as labels, the input feature vectors and labels were used by SVM and random forest algorithms, and binary classifiers were output for direct symbol detection. In re-processing, this work first estimated the ambient RF signal by an MMSE estimator and then eliminated it from the original received signal to enhance the power level of useful backscatter signals. It is worth noting that such pre-processing is a software DLI cancellation. The error rates of the proposed methods were lower than ED and traditional MMSE detection, especially under low SNR.

The KNN method was proposed in [17]. In this study, an AmBC system with a ULA antenna array receiver and adopting BPSK modulation, as well as the protocol in Fig. 2(B), was considered. The samples of training pilot signals were used as input feature vectors directly to calculate distances to their k nearest neighbors. Besides, beamforming technology was utilized to distinguish DLI and backscatter signals, and Hadamard encoding was exploited to correct weak signal transmission. In pre-processing, the receiver estimated the angle of arrival (AoA) of the DLI first to obtain the weight vector. Then according to the weight vector, the useful backscatter signal was shifted to an orthogonal signal space. After that, the samples of clear backscatter signals were regarded as the input vectors of the KNN algorithm, and the output was the classifier. Since the symbols were encoded by Hadamard codeword, the receiver can decode it for error correction, which can promote detection accuracy. The proposed method can achieve non-error receiving in the high SNR region when the distance is not too long.

The UL method based on the energy set and the EM algorithm was introduced in [18]. It studied a typical AmBC system, assuming that the ambient RF source adopted an equal amplitude modulation, and so the average energy levels of the received signal only depended on load modulation at BDs. However, the difference between energy levels was generally very small because of strong DLI. To avoid directly using energy as an input feature, the authors built mixture distributions to describe energy features when BDs transmitted different symbols. Then, the corresponding distribution parameters of different symbols were estimated by using the EM algorithm. This method has good performance when the BD is not far away from the receiver, but when the distance increases or the channel states change, the clustering accuracy will decline.

The above methods are all based on energy feature engineering, which is limited by DLI. To expand alternative features, the authors in [19] explored a UL method based on the pattern of the signal constellation. The authors found that the received superimposed signal can maintain the constellation pattern similar to the ambient RF source signal, and thus they considered constellation pattern as a feature to distinguish different symbols. This work first assumed the received superimposed signal followed the Gaussian Mixture model (GMM) and mined the corresponding relationship between the received signal and constellation patterns. The clustering results were solved by the EM algorithm. Then different symbols were matched with clustering results according to labels. In the above process, DLI did not affect the generation of constellation patterns, so the performance of the proposed method can overcome the technical challenges faced by energy features and achieve better performance than energy-based feature engineering.

DL Methods Based on ANN: The ML method based on feature engineering relies too much on background knowledge and feature engineering experience. Limited by that, the designed training model is not always optimal. Instead, the DL methods based on ANN can extract and analyze the hidden features of the raw input data through the designed neural networks. DL methods have made many achievements in wide fields, e.g., traditional wireless communication, natural language processing, computer vision, etc. In AmBC signal detection, the DL methods are expected to combine with the multi-antenna receiver to explore abundant instinct features (e.g., temporal and spatial features) in raw data to achieve better detection performance [23].

The article [20] designed a deep transfer learning framework suitable for AmBC signal detection. It considered a classical AmBC system using the multi-antenna receiver and adopted the classical communication protocol. The symbol detection problem using OOK modulation was transferred into a binary hypothesis detection problem in this research. To utilize rich temporal and spatial features, the covariance matrix was used as input data to extract features via two convolution layers. The parameters obtained after training can obtain the optimal likelihood ratio test (LRT) performance, which is close to the optimal signal detector with perfect CSI.

4 Future Research Trends

4.1 Enriching Data Sets

At present, the data sets used in ML methods of AmBC channel estimation and signal detection are mainly manually simulated based on the typical three-node AmBC system. There is still a lack of real data in field experiments of actual communication scenarios, let alone standard data sets for scholars to conduct extensive research on ML methods. It is of great significance to enrich AmBC data set generation methods and construct real data sets for AmBC channel estimation and AmBC signal detection.

4.2 Improving Model Accuracy

Both supervised learning and unsupervised learning require a large number of training data to ensure the accuracy of the trained model. However, online training with typical communication protocols cannot provide such a large scale of data. The current research transmits known content pilot symbols to obtain the original training data set with labels for offline training. However, the environment is volatile, and the offline data sets usually cannot accurately describe the real-time channels, and thus the accuracy of the ML method is reduced. To solve this problem, transfer learning can conduct abundant offline training to obtain the basic parameters of AmBC. Few fine-tuning steps are conducted to make the trained model more accurate for real-time channels in online training. It accelerates the training speed in new scenarios. In addition, data enhancement techniques can be adopted to maximize the utilization of limited pilot symbols.

4.3 Reducing Training Costs

The computational complexity of ML methods is generally higher than traditional methods. Both supervised learning and unsupervised learning methods usually require a large of computational resources for model training. Meanwhile, deep learning methods have higher computation requirements because they use multi-layer neural networks. The computational cost of these methods may not be afforded in many AmBC systems. Simplifying training models or neural networks and reducing the training overhead of ML methods are crucial in ultra-low-power IoT scenarios.

4.4 Expanding Application Scopes

The ML methods discussed in this paper are mainly used in the physical layer to handle channel estimation and signal detection challenges in AmBC systems. In the future, AmBC is supposed to be applied in large-scale IoT networks, where it will face more challenges in the MAC layer and network layer. Although ML methods have been used in traditional wireless communications from the physical layer to the application layer, the unique technical challenges discussed in AmBC (e.g., energy constraints, noncooperation of ambient RF source, and lack of perfect CSI) are still major obstacles to ML methods in other layers of AmBC networks. Expanding the application scopes of the ML method AmBC still needs to solve these technical challenges.

5 Conclusion

In this article, we provide a survey on ML-based methods for AmBC channel estimation and signal detection. We first provide a brief overview of AmBC from its brief history, architecture, and working paradigms. Then we summarize signal-receiving issues in AmBC, where we discuss the challenges in different aspects and highlight signal processing key technologies. After that, we focus on machine learning methods for AmBC channel estimation and AmBC signal detection, including the goals, technical challenges, classical solutions, and current ML-based methods. Finally, we introduce many future research trends.

Acknowledgements. This work was supported in part by the Beijing Natural Science Foundation (Grant No. JQ21036), the National Natural Science Foundation of China (Grant No. 62293494, No. 62301078, No. 61821001, No. 62271086), the China Postdoctoral Science Foundation (Grant No. GZB20230086), and the Beijing Key Laboratory of Work SafetyIntelligent Monitoring.

References

1. Van Huynh, N., Hoang, D.T., Lu, X., Niyato, D., Wang, P., Kim, D.I.: Ambient backscatter communications: a contemporary survey. IEEE Commun. Surv. Tutor. **20**(4), 2889–2922 (2018)
2. Ma, S., Wang, G., Fan, R., Tellambura, C.: Blind channel estimation for ambient backscatter communication systems. IEEE Commun. Lett. **22**(6), 1296–1299 (2018)
3. Altuwairgi, K.H., Tota Khel, A.M., Hamdi, K.A.: Energy detection for reflecting surfaces-aided ambient backscatter communications. IEEE Trans. Green Commun. Netw. **8**(1), 279–290 (2024)
4. Wang, G., Gao, F., Fan, R., Tellambura, C.: Ambient backscatter communication systems: detection and performance analysis. IEEE Trans. Commun. **64**(11), 4836–4846 (2016)
5. Jing, F., Zhang, H., Gao, M., Xue, B.: Bayesian-mle signal detection for multi-antenna ambient backscatter communication. IET Commun. **16**(6), 672–684 (2022)
6. Sun, Y., Peng, M., Zhou, Y., Huang, Y., Mao, S.: Application of machine learning in wireless networks: Key techniques and open issues. IEEE Commun. Surv. Tutor. **21**(4), 3072–3108 (2019)
7. Liu, V., Parks, A., Talla, V., Gollakota, S., Wetherall, D., Smith, J.R.: Ambient backscatter: wireless communication out of thin air. ACM SIGCOMM Comput. Commun. Rev. **43**(4), 39–50 (2013)
8. Wang, A., Iyer, V., Talla, V., Smith, J.R., Gollakota, S.: FM backscatter: enabling connected cities and smart fabrics. In: 14th USENIX Symposium on Networked Systems Design and Implementation (NSDI 17), pp. 243–258 (2017)
9. Kellogg, B., Parks, A., Gollakota, S., Smith, J.R., Wetherall, D.: Wi-Fi Backscatter: Internet Connectivity for RF-Powered Devices, pp. 607–618 (2014)
10. Ensworth, J.F., Reynolds, M.S.: Every smart phone is a backscatter reader: modulated backscatter compatibility with Bluetooth 4.0 low energy (BLE) devices. In: 2015 IEEE International Conference on RFID (RFID), pp. 78–85. IEEE (2015)
11. Peng, Y., et al.: PLoRa: a passive long-range data network from ambient lora transmissions. In: Proceedings of the 2018 Conference of the ACM Special Interest Group on Data Communication, pp. 147–160 (2018)
12. Griffin, J.D., Durgin, G.D.: Link envelope correlation in the backscatter channel. IEEE Commun. Lett. **11**(9), 735–737 (2007)
13. Ma, S., Zhu, Y., Wang, G., He, R.: Machine learning aided channel estimation for ambient backscatter communication systems. In: 2018 IEEE International Conference on Communication Systems (ICCS), pp. 67–71. IEEE (2018)
14. Abdallah, S., Verboven, Z., Saad, M., Albreem, M.A.: Channel estimation for full-duplex multi-antenna ambient backscatter communication systems. IEEE Trans. Commun. (2023)
15. Liu, X., Liu, C., Li, Y., Vucetic, B., Ng, D.W.K.: Deep residual learning-assisted channel estimation in ambient backscatter communications. IEEE Wirel. Commun. Lett. **10**(2), 339–343 (2020)
16. Hu, Y., Wang, P., Lin, Z., Ding, M., Liang, Y.-C.: Machine learning based signal detection for ambient backscatter communications. In: ICC 2019-2019 IEEE International Conference on Communications (ICC), pp. 1–6. IEEE (2019)
17. Wang, X., Duan, R., Yigitler, H., Menta, E., Jantti, R.: Machine learning-assisted detection for BPSK-modulated ambient backscatter communication systems. In: 2019 IEEE Global Communications Conference (GLOBECOM), pp. 1–6. IEEE (2019)

18. Zhang, Q., Liang, Y.-C.: Signal detection for ambient backscatter communications using unsupervised learning. In: 2017 IEEE Globecom Workshops (GC Wkshps), pp. 1–6. IEEE (2017)
19. Zhang, Q., Guo, H., Liang, Y.-C., Yuan, X.: Constellation learning-based signal detection for ambient backscatter communication systems. IEEE J. Sel. Areas Commun. **37**(2), 452–463 (2018)
20. Liu, C., Wei, Z., Ng, D.W.K., Yuan, J., Liang, Y.-C.: Deep transfer learning for signal detection in ambient backscatter communications. IEEE Trans. Wirel. Commun. **20**(3), 1624–1638 (2020)
21. Abdallah, S., Verboven, Z., Saad, M., Albreem, M.A.: Channel estimation for full-duplex multi-antenna ambient backscatter communication systems. IEEE Trans. Commun. (2023)
22. Liu, C., Liu, X., Ng, D.W.K., Yuan, J.: Deep residual learning for channel estimation in intelligent reflecting surface-assisted multi-user communications. IEEE Trans. Wirel. Commun. **21**(2), 898–912 (2021)
23. Guo, S., Zhao, X., Zhang, W.: Throughput maximization for RF powered cognitive Noma networks with backscatter communication by deep reinforcement learning. IEEE Trans. Wirel. Commun. (2023)