

**Titel der Arbeit**

Improvement of RFID-based electronic article surveillance performance using machine learning methods

**Management Summary: Kurzfassung der Arbeit (max. 3000 Zeichen)**

Diebstahl ist die häufigste Ursache von Schwund im Inventar im Handel. Hinzu kommt, dass herkömmliche Diebstahlschutz-Systeme sowohl in der Anschaffung als auch im Betrieb teuer sind. Vor diesem Hintergrund wird in der Bachelorarbeit „Improvement of RFID-based electronic article surveillance performance using machine learning methods“ der Einsatz von RFID zur elektronischen Diebstahlsicherung im Einzelhandel betrachtet. Wenn Unternehmen RFID bereits in der Supply Chain einsetzen, entstehen keine weiteren Kosten für das Ausrüsten der Produkte mit RFID-Tags. Bei herkömmlichen Systemen sind die Tags meist weit teurer als RFID-Tags und müssen vom Personal zusätzlich angebracht und nach dem Verkauf entfernt werden. Diese Tags können durch RFID abgelöst und so die Kosten für die Diebstahlsicherung gesenkt werden. Zudem können so Verluste durch Diebstahl verhindert werden. Auch für den Fall, dass ein Ladendieb entkommen kann, bietet RFID Vorteile gegenüber den herkömmlichen Systemen: Die gestohlenen Artikel können dank des auf den Tags gespeicherten EPC identifiziert werden und daraufhin das Bestandsmanagement benachrichtigt werden.

Bei RFID-Datenströmen handelt es sich um stark verrauschte Datenströme, wodurch für die RFID-basierte Diebstahlsicherung bislang keine ausreichend hohe Erkennungsgüte erreicht werden konnte. Das Ziel der Arbeit ist die Entwicklung eines vollautomatischen Verfahrens zur Unterscheidung von RFID-getaggtter Ware, die eine Filiale durch ein Diebstahlsicherungsgate verlässt und sonstigen RFID-getaggtten Artikeln, die rein zufällig im Lesebereich der Antennen gelagert oder an ihnen vorbeigetragen werden. Auf physikalischer oder elektrotechnischer Ebene konnte diese Unterscheidung bislang nicht mit ausreichender Güte erreicht werden. Aus diesem Grund wurde in der Arbeit untersucht, inwieweit mit Methoden des Data Mining zwischen den beiden Fällen unterschieden werden kann. Hierfür wurden im RFID-Labor der Universität Parma Experimente durchgeführt, die typischen Laufwegen von Kunden im Handel mit getaggtten Waren entsprechen. Es wurden die Rohdaten der Antennen vorverarbeitet und anschließend in maschinelle Lernverfahren eingespeist, die erst dann Alarm schlagen sollen, wenn der mutmaßliche Dieb das Gate am Ladenausgang durchläuft. Hierfür wurden über 1.200 Testdurchläufe durchgeführt und mehr als 150.000 Lesungen aufgezeichnet. Danach wurden auf den Rohdaten für jeden Testdurchlauf mehr als 40 Prädiktoren entwickelt. Mithilfe dieser Prädiktoren wurden im Anschluss verschiedene Data-Mining-Modelle trainiert. Mit diesen Modellen konnte in über 99% der Testdurchläufe richtig zwischen „Diebstahl“ und „kein Diebstahl“ unterschieden werden. Die besten Ergebnisse lieferten mit dem Neuronalen Netz und dem Random Forest sogenannte Black-Box-Verfahren. Data-Mining-Modelle eignen sich folglich sehr gut zur Verbesserung der RFID-basierten elektronischen Diebstahlsicherung.

**Stellungnahme zu den folgenden bewertungsrelevanten Aspekten der Arbeit**

**1. Relevanz für den Handel: Ist das Thema bereits im Handel erforscht worden? Welche neuen Erkenntnisse für den Handel hält die Arbeit bereit? (Max. 1250 Zeichen)**

RFID wird bei vielen Handelsunternehmen entlang der Supply Chain zur Warenidentifikation und Erfassung von Warenströmen eingesetzt. Diese Prozesse zeichnen sich in der Regel durch eine hohe Kontrollierbarkeit aus (z.B. Einweisung Personal durch das Unternehmen). Die Bachelorarbeit beschäftigt sich hingegen mit einem Prozess in der sogenannten „letzten Meile“ der Supply Chain mit im Vergleich zu vorgelagerten Prozessen deutlich eingeschränkter Prozesskontrollierbarkeit. Die Artikel befinden sich hierbei in den Händen der Kunden, welchen ein Unternehmen natürlich – im Gegensatz zu den eigenen Angestellten – keine Vorschriften im Umgang mit der Technologie machen kann. Diese Fehlerquellen (z.B. Abschirmungen oder Reflexionen) führen dazu, dass RFID heute trotz des großen Potentials der Technologie nur vereinzelt zur Diebstahlsicherung eingesetzt wird. In der Arbeit wird ein Ansatz vorgestellt, mit dem eine hohe Prozessqualität für die RFID-basierte Diebstahlsicherung erreicht werden kann. Dadurch wird es möglich, die Vorteile der Technologie für den Handel zu heben.

**2. Umsetzbarkeit im Handel: Wo liegt der Anwendungsnutzen für den Handel? (Max. 1250 Zeichen)**

Bei Nutzung von RFID zur Diebstahlsicherung fallen keine weiteren Tagging-Kosten an, sofern das (Handels-)Unternehmen RFID bereits in der Supply Chain einsetzt. Teure proprietäre Systeme, bei denen die Tags nicht selten 0,50€ pro Stück kosten und zusätzlich noch vom Personal händisch angebracht und entfernt werden müssen, können so abgelöst werden. Zudem können Artikel mithilfe von RFID-Tags identifiziert werden. Dadurch wird die Bestandsgenauigkeit deutlich erhöht, da die Warenbewegung erfasst wird, selbst wenn ein Dieb entkommen sollte. Außerdem werden durch die hohe Erkennungsrate die Verluste durch Diebstahl – der größten Quelle von Schwund im Inventar – verringert, während gleichzeitig die Kundenzufriedenheit dank der minimalen Rate an Fehlalarmen nicht gefährdet wird.

### 3. Innovationskraft der Idee (Max. 1250 Zeichen)

Die Arbeit ist die erste wissenschaftliche Arbeit, die mit Data-Mining Modellen den Prozess der RFID-basierten Diebstahlsicherung verbessert. Dafür werden die von den Antennen registrierten Signalstärken der einzelnen Lesungen in zeitliche und räumliche Relation zueinander gesetzt und die dadurch gewonnenen Informationen zum Training der Modelle genutzt. In früheren Arbeiten, z.B. in einer 2011 veröffentlichten Studie von Prof. Bottani, wurden für RFID deutlich schlechtere Diebstahl Erkennungsraten als für herkömmliche Technologien festgestellt. Einer der Gründe hierfür ist, dass – analog zu den anderen eingesetzten Technologien – ein Alarm ausgelöst wurde, sobald die RFID-Antennen des Sicherungsgates die Lesung eines Tags eines nicht verkauften Artikels registrieren. Um die Lesung von Kleidungsstücken zu verhindern, die über einen Meter vom Gate entfernt sind (z.B. Ausstellungsstücke), wurde die Sendeleistung der Antennen stark gedrosselt. Diese Drosselung bedingt die niedrigen Erkennungsraten. Eine höhere Sendeleistung hingegen – welche die Erkennungsraten verbessern würde – erhöht die Anzahl der Fehlalarme, da sich dadurch das Lesefeld der Antennen vergrößert. Mit Data-Mining Modellen kann dieses Dilemma aufgelöst werden.

### 4. Angewandte wissenschaftliche Methode (max. 600 Zeichen)

Struktur und Aufbau orientieren sich grob am CRISP-DM Prozess, dem De-Facto-Industriestandard für Data-Mining Projekte. CRISP-DM umfasst die Phasen Business Understanding, Data Understanding, Data Preparation, Modeling sowie Evaluation. So wird zunächst die Problemstellung beschrieben und der Nutzen aus Sicht des Handels ermittelt. Anschließend werden die Versuchsreihen erläutert, die Daten beschrieben und analysiert, aufwendig vorverarbeitet und anschließend in aggregierter Form zum Training der Data-Mining Modelle genutzt. Im letzten Schritt werden die Ergebnisse dargestellt und bewertet.