On the Superdistribution of Digital Goods

Andreas U. Schmidt
(CREATE-NET Research Consortium
Via alla Cascata 56/D, 38100 Trento, Italy
andreas.schmidt@create-net.org)

Abstract: Business models involving buyers of digital goods in the distribution process are called superdistribution schemes. We review the state-of-the-art of research and application of superdistribution and propose a systematic approach to market mechanisms using superdistribution and technical system architectures supporting it. The limiting conditions on such markets are of economic, legal, technical, and psychological nature. Large scale applications of superdistribution such as video-on-demand and multimedia over peer-to-peer type networks pose particular requirements on security and efficiency.

Keywords: Content superdistribution, digital good, copyright protection, trusted computing

Categories: C.2.0, H.1.1, K.4.2, K.5.1

1 Introduction

Information systems in general and the distribution of digital content in particular are dominated by centralised structures rooted in client-server models, and large efforts have been made for the vertical integration of content production, ingestion, and distribution [Axmedis 09]. The final transportation of content to the head-ends is nowadays either digital broadcast, e.g., DVB [Reimers 06], multicast, as for instance envisioned in 3GPP Long-Term Evolution [3GPP 08], or content push [OMA 06].

Peer-to-peer (p2p) systems on the other hand realise a completely different paradigm for data transport in networks, namely distribution from nodes to other nodes with little involvement of central instances [Androutsellis-Theotokis and Spinellis 04]. File-sharing networks like KaZaA or Gnutella embody this paradigm on the application level, implementing overlay networks in which users actively (with varied degrees of automation) re- or superdistribute content, in the form of digital files, to other users.

The term superdistribution may have been coined in [Mori and Kawahara 90, 97], in any case it has been around in information and communication research for some time. Though the concept lay dormant for quite a while — perhaps due to the association with the dominant use of p2p and file sharing by free riders and the copyright wars — interest in superdistribution has been rekindled recently in the content producing industry. The combined size of the most important existing businesses based on content superdistribution schemes are of a small scale in comparison to the turnovers of the media industry as a whole. Nevertheless they prove that the industry is seriously experimenting with the concept. Most importantly,
superdistribution has even been cast in the form of a standard for the mobile domain by the Open Mobile Alliance (OMA).

Technically, superdistribution has hitherto been viewed just as a variant of Digital Rights Management (DRM) [Becker et al. 03], [May (07)], or of p2p systems, and research on its fundamentals is still scarce. For instance, basic economic questions pertaining to the viability of superdistribution in particular in competition with free riders have only been examined in our previous work [Schmidt 08a].

The present paper presents is a survey contributing a first treatment of characteristic issues of superdistribution systems differentiating them from DRM, viewed as information systems in their application and economic context. The line of argument is as follows. A system model for generic superdistribution is proposed in [Section 2] and used throughout the paper. [Section 3] gives an overview over “historic” and current superdistribution systems which have been deployed in the real world. From these examples and other research sources we derive in [Section 4] central systemic, non-technical requirements on successful superdistribution systems. These insights are the conceptual background behind the inception of the multimedia superdistribution system of the NanoDataCenters EU project described in [Section 5]. In [Section 5.1] the high-level demands of [Section 4] are mapped to privacy, security, and functional requirements in the special context of multimedia superdistribution with devices in users’ homes. [Sections 5.2 and 5.3] propose to use Trusted Computing as a core technology to implement these requirements. [Sections 5.4 and 5.5] elucidate the trust relations between different stakeholders in the content distribution network based on the “trusted set-top boxes” and outline business models in this context. The example of [Section 5] is presented on a par with those of [Section 3] with no more technical detail. The latter is deferred to forthcoming publications [Bal, Kuntze, Schmidt 09], [Brett, Kuntze and Schmidt 09]. [Section 6] closes the loop by putting superdistribution into the context of current socio-economic developments surrounding content distribution, copyright protection, and piracy, in front of their historic background. We conclude in [Section 7].

2 The General Structure of Superdistribution Networks

Superdistribution is the combined distribution and market scheme for digital goods involving buyers in the distribution process in such a way that they redistribute the good to other legitimate buyers. Here a digital good is an information good in the economical sense [Shapiro and Varian (99)], [Stegman 04], which is represented in digital form, regardless of being embodied physically or only in intangible form (some use the term virtual goods, coined by [Aichroth and Hasselbach 03] and used for information goods in intangible, digital form, and distributed via electronic networks). In an active sense, to superdistribute means the combined transaction of acquiring a good and its (offering for) re-distribution, or resale, and actually transferring it to another node.

Here we argue that existing system models for DRM are too narrow to accommodate for the specific features and structures of superdistribution. In fact, extending DRM into various directions is a recent research trend, which is triggered by the manifold ways in which users operate with digital goods for instance in social networks. For instance, [Stini, Mauve, and Fitzek 06] transcend DRM by envisioning
a system in which only the information on “who owns this digital good” is managed and thus agents in the economic network can be given ample freedom, e.g., to superdistribute it. In this section, we present a similar approach to extend information management systems in an appropriate way for superdistribution.

2.1 Superdistribution networks

A superdistribution network in the most general sense has two sides. The first one is the network over which the good is distributed, economically a logistics network for the final distribution of the good to the consumer. If it is an electronic network, it is a particular kind of a content distribution network, like Akamai, Amazon S3, Corel, CDNetworks, etc. Whether a good is distributed by ordinary mail, over an electronic network, or by short-range communication between mobile devices, is immaterial for the classification as a superdistribution network. Paradigm examples exist for all three variants: Superdistribution by mail is represented by the classic chain letter, peer-to-peer networks are the paradigm for superdistribution over electronic networks, and superdistribution between mobile devices is for instance standardised by OMA. In all cases, superdistribution is an overlay over an (often general-purpose) communication or transportation network, like ordinary mail, the Internet, or Bluetooth ad hoc communication between mobile devices. We call this side of a superdistribution network the content distribution overlay (CDO). The CDO is a directed graph, which in most (reasonable) cases may be assumed to be a connected tree. The CDO graph can be coloured, i.e., various attributes may be attached to the edges, a particular example being that the quality of the good may change, e.g., improved by a superdistributing node to compete with other resellers.

Figure 1: Superdistribution overlay networks in the system context
Superdistributing nodes in a CDO need a good, economic reason to participate. This is always true due to the minimum marginal cost greater than zero incurred by a superdistributing node for storage and transferral of the good to and from him-/herself (one of the two at least is borne by a specific node). That is, nodes expect some kind of remuneration for participating actively in the CDO — otherwise they may just become sinks for the digital good. The flow of remuneration — pecuniary, informational, immaterial, or of any other conceivable kind — constitutes another overlay network, the remuneration overlay network (RON). The claim here is that no superdistribution network exists without RON, the most trivial example being the tree spanned by the resale prices paid by buyers to superdistributing nodes in the CDO. In this case the edges of the RON are just the edges of the CDO with inverted directions (and different colours, e.g., the sales price, attached). The node-set of the RON can be assumed to be a subset of the node-set of the CDO, but the relation of the RON’s edges to the edges of the CDO is generally nontrivial. For instance in multi-level marketing, a buyer of a good might pay a reward to resellers further down the line, and not only the resale price to his direct reseller. [Fig. 1] shows CDO and RON in the context of underlying communication and payment networks.

2.2 Digital goods

The term digital good used so far refers to the economical atom distributed over the CDO and being the root cause for the RON. Informationally, the digital good is a compound minimally consisting of three components. The content is the piece of digital information that is actually used and, if the node chooses to do so, offered for distribution to others. As the superdistribution network is an economic market mechanism, the content is necessarily accompanied by information representing the contractual rules of a) the global superdistribution market, and b) the particular relationship between superdistributing (reseller) and acquiring (buyer) node. Though we will not make use of this distinction of local and global contract, this orthogonal categorisation may be useful, e.g., to classify superdistribution networks.

Using the good means, on the one hand, that the content is consumed by a node who acquires it. Consumption of the content represents one part of the value proposition that the digital good represents to a buyer. It is governed by a piece of information commonly called the consumption licence, which describes the conditions and permissions under which the buyer can use the content. Economically speaking, the consumption licence prescribes the ways in which a buyer may turn the value proposition of the content into utility. The consumption licence is also thought to be the informational link between the digital good and the remuneration overlay by stating the rules of payment for the good to the superdistributing node, as well as any other reward to be paid to further nodes or entities. In this way the consumption licence generates the RON from the CDO, assuming the rules are adhered to by all participants. Summarising, the consumption licence consists of three parts:

- Consumption rules describe how content may be used;
- Remuneration rules describe how and who must be paid for it;
- The Content association describes to which content the rules apply.

The second way in which an acquiring node can make use of the good is by superdistribution. We think of it as governed by rules incorporated in a second licence, the redistribution licence. Just as the consumption licence connects the good
to the RON, the redistribution licence conditions, or generates the content distribution overlay. Thus, this licence consists of two essential parts:

- **Redistribution rules** describe how, to whom and under which conditions content may be redistributed;
- The **Content association** describes to which content the rules apply.

The complete informational structure and its relation to CDO and RON is visualised in [Fig. 2].

![Figure 2: Information model of superdistribution.](image)

Of course, many other groupings of the information characterising a superdistribution network are possible — the approach chosen here is lead by the distinction between CDO and RON. It should also be noted that all notions introduced above are understood here in the broadest possible sense. That is large parts of the rules and licences may be represented differently than in digital form and may include for instance general legislation, copyright law, social norms, etc. Redistribution in particular can also be governed by technical conditions, e.g., the information system that represents the platform for the execution of superdistribution.

Thus the particular rules that need to be represented digitally in a concrete superdistribution network may be restrictions as well as extensions of such global, or external, rules. Likewise, content associations may be simple titles, digital identifiers denoting a single piece of content or a group, or be augmented by information protecting the integrity of digital content such as hash values or signatures. Nothing restricts the methods by which the licences and the content are generated, stored, and transferred in the superdistribution network. This conceptual approach is well known from general DRM [Becker et al. 03].

### 2.3 Examples

Some more concrete examples might elucidate the abstract notions of Section 2.2. The most direct form of remuneration is a resale price paid by the acquiring node to the superdistributing node. This makes the superdistribution network a genuine network market of buyers/resellers, where an incentive to buy a good accrues to them
by the resale revenues they can achieve. The “multi” in the term multi-level marketing often refers to the fact that many subsequent levels or generations of buyers contribute to a node’s resales revenues, or even all of them. This kind of payments or remunerations from the downline may be restricted to a finite number of buyer generations or not, the latter case being realised in some network marketing schemes for physical goods.

The remunerations may be conditioned by various global or individual factors such as time, buyer/reseller location, distance in a social network, or externalities like a measured popularity of the content. In many cases it makes sense to let a part of the resales price accrue to a central entity external to the CDO proper, which we call collector. Its role may be to skim revenues from the market for, e.g., the artists and or labels, or it may act as a (state) collecting agency implementing taxation on the distribution of digital goods. Second-level payments are represented by the thick dashed arrows in [Fig. 1]. The collector is shown there as an external entity directly remunerating the Originator. In reality this might involve a payment provider or network operator as well.

An interesting example for restrictions on the redistribution is the implementation of territorial protection. This can be used to protect resellers from the competition of their (direct) buyers to a certain extent buy stating, e.g., “do not superdistribute before moving away by 100 metres”. Thus, this kind of redistribution rule using restrictions based on geographical location may make particular sense for CDO based short-range communication between mobile users, i.e., mobile superdistribution. We showed in [Schmidt, Kuntze, and Abendroth 08] how such conditions can be enforced in an efficient, decentralised, yet secure manner.

3 Some Examples

As said, superdistribution networks occupy only a small niche even of the online content distribution market. The better known examples are the following. Snocap [Snocap 08], founded by one of the fathers of Napster, was started with the idea to obtain licences from the music industry which explicitly allow to distribute content over the existing, popular p2p networks. Snocap uses audio fingerprinting to track the distribution of content, and file-sharing networks need to be adapted to support Snocap’s remuneration scheme. Though Snocap has made some deals with many, even major, labels, it never took off economically and the company has been acquired in February 2008 by the social networking platform imeem [imeem 08]. After restructuring and changing the strategy, Snocap has become a general service provider for online music distribution and for instance provides the technology for the music stores in MySpace. MashBoxx [MashBoxx 08] started with similar ambitions and also close to the circles of Napster and Grokster, see [Menn (03)], [EFF Grokster 08]. The company seems to lay dormant for some time, appearing in the news only for recent intellectual property litigations. Peer Impact is a pay-for-download file-sharing service created by Wurld Media, and now acquired together with its parent company by the online video service provider Roo, see [Roo 08]. The file-sharing client has been re-released under the new brand name ToPeer [2peer 08], which seems to use part of the original technology to allow p2p users to create private spaces in which to share content with peers they trust.
The Paradiso system [Paradiso 08], [Nair, Gerrits, Crispo, and Tannenbaum 08] is a technological solution to DRM-based superdistribution with strong security properties. Its central technical trait is that it relies on a trusted platform [Fichtinger et al. 08] to ensure adherence to consumption and redistribution licence. Thus it poses technical requirements on compliant devices with regard to cryptographic capabilities (hash, AES engine, and PKI management), secure storage, and secure content decoder. In the content distribution scheme of Paradiso, consumption and remuneration licences are cryptographically bound to the content and chained. That is, a buyer receives with the content a signed container from the reseller, containing all previous licences created in every resale upstream in the CDO. The signature also associates this data to the content. This enables him, e.g., to verify that the content has not been tampered with, for instance it prevents content masquerading attacks by which a reseller might try to superdistribute content of lower quality. The compliant device can also check that all licence rules have been enforced in all previous distribution steps, and enforces the applicable rules for itself, e.g., respects and updates the allowed number of resales. Payment is an out of band process in Paradiso which is based on a receipt the acquiring node sends back to the superdistributing node. It is not hard to see that this system has strong security with respect to the maintenance of DRM of the content as it is distributed down the CDO. Formal security proofs are given in [Jonker, Nair, and Dashti 06]. This system provides the strongest possible DRM enforcement in superdistribution which can be implemented in a completely decentralised fashion. In the following we describe a different example in more detail.

3.1 Potato system

The Potato system [Potato 08] is a product developed by the 4FO AG [4FriendsOnly 08] (founded in 2000) together with the Fraunhofer Institute for Digital Media Technology IDMT [IDMT 08] in Ilmenau, Germany, for superdistribution of music as mp3-files. The technical platform for superdistribution presented by the Potato system is centralised, insofar as it uses a central accounting service (AS) for registration and publishing new songs by originators, and to operate the remuneration scheme. The content CDO is completely free of any DRM measure. The only information protected by the AS (besides the content integrity of which is proved by a hash value) is the redistribution licence, which is obtained by a buyer upon payment in the form of a transaction number (TAN). The TAN serves as a receipt which is simply added to the file name, which is in turn announced in subsequent resales to the AS which initiates the rewarding of resellers. Some details are found in [Grimm and Nützel 02], [Nützel and Grimm 03]. Potato supports various payment providers from which the originators of a good may choose.
The market mechanism and remuneration scheme implemented in the Potato system is perhaps the most evolved in superdistribution. The sharing of revenues is shown in [Fig. 3]. Potato targets small labels and independent artists, who may obtain 55–70% of the purchase price of every resale, depending on the service level they choose. An interesting detail is that Potato has an agreement with the German collecting society for music, GEMA which obtains the due contributions directly from the system. Potato itself and the payment provider share 14% of the purchase price and further 14% are distributed as resale revenues from the buyer to resellers (this share has been decreased from 35% in “Version 1.0” to the current “Version 2.0” value). The special kind of remuneration for resellers in this system establishes a true multi-level market with three rewarding levels, each being awarded a geometrically decreasing share of 10, 3, and 1%, respectively; cf. [Schmidt 08a, Section II.B]. Thus the CDO and RON look locally as shown in [Fig. 4]. It is interesting to note that a rebate of 2% (borne by the system, not the resellers) is offered for nodes who choose to buy from a peer rather than the central service. This is an important incentive that reduces the dominant role of a single market participant, cf. [Section 4.2].

Originally resellers were mostly left to their own devices in marketing songs for resale. They could use a resale link containing their TAN on their Web-site or in e-mails. The most recent developments of the Potato as a superdistribution platform
regard capabilities to support users in marketing goods, i.e., means to offer them successfully online for superdistribution, and to compete with other resellers. This includes the extension of resale links to Widgets embodying small online shops where resellers can display their favourites, covers, and let peers listen to clippings of songs. 4FO also added a social commerce platform SpreadBox [SpreadBox 08] to its portfolio which also tries to leverage community aspects of marketing in the form of product recommendation.

4 Conditions for viable superdistribution

Superdistribution may seem a variant of DRM or p2p, or a combination of both. Now we try to elaborate on specific traits to show how superdistribution is different.

4.1 A. The axis of lawfulness and legitimacy

The RON, if effective, turns superdistribution into a network marketing scheme, or if multiple buyer generations receive remunerations, a multi-level marketing scheme. Multi-level marketing carries negative connotations and is illegal in special forms known as pyramid selling, snowball systems, chain letters, etc., under many jurisdictions. This similarity to illicit schemes has perhaps also impeded applied research in the field of superdistribution as such. The authors of [Micklitz, Monazzahian, and Rößler 99, Vol. II] present criteria to distinguish between legitimate multi-level marketing and such practises that are to be considered illicit. In the case of digital goods some arguments speak for the viability of fair superdistribution schemes (thoroughly discussed in [Schmidt 05], [Schmidt 06]). i) Buyers acquire not only a void right to resale, but also a good of value. Potential losses an agent entering at a late stage will incur are charged up against this value; ii) Inventory loading, i.e., the obligation to keep a large, non-returnable stock, is irrelevant for digital goods; iii) Marginal costs for replication and redistribution are mostly much smaller than resale prices and thus transaction costs are largely insignificant; iv) A main novel feature of the concepts above is that they enable in principle a fair system design see [Section 4.2].

Other legal requirements for superdistribution are derived from the corresponding ones for general electronic commerce. i) privacy of buyers and sellers should be maintained by implementing minimal-need-to-know principles; ii) Consumer protection legislation, as, e.g., in the EU [EC Consumer Law 08], needs to be respected; iii) Copyright law must be respected, i.e., originators rights must be properly transcribed into the licences and a system’s operator must obtain all necessary rights and involve collecting societies, etc. iv) Contracts between buyers and resellers must be enforceable and individual fraud (e.g., by selling content of lower value than proposed) must be prevented; v) Market abuse and distortion must be prevented, cf. the economical and security requirements below.

4.2 The economical axis

Digital goods share the properties of information goods which are transferable and non-rival like public goods, and additionally are durable, i.e., show no wear out by
usage or time [11], [12]. Like for a private good, however, original creation can be costly, whereas reproduction and redistribution are potentially very cheap. This is the economical basis for superdistribution which emulates the distribution system of free-riders, namely p2p networks [Zerfiridis and Karatza, 04]. They pose additional value proposition to buyers of the original (legal) version of the good by revenues or other rewards linked to resales. Thus the central question for superdistribution of digital goods is economic viability in the presence of free-riders.

The RON of a superdistribution network is a network marketing scheme. Theoretical treatments for network markets are scarce, which inspired us to devise a stochastic model for the dynamics of such markets in [34] and evaluate it in various ways [10], [35]. The model is essentially comprised of atomic agents entering the market continuously until saturation, with equal chance to trade with each other, i.e., to buy the good from a reseller. With these assumptions, a node entering the CDO at a certain point in time, i.e., a certain market saturation, can calculate its expected revenues from subsequent resales, given that the price schedule of current and future resales prices is known. [Fig. 5] shows two examples (black, blue) of prices (dashed), expected resales revenues (thin solid) and effective prices, i.e., price paid minus expected revenues (thick solid), plotted against the saturation parameter running from 0 to 1. The thrilling flip side of the innocuous mathematical expressions defining this model is that it enables dynamical forward pricing. That is, the operator of a superdistribution network can in principle control the incentive that accrues to buyers via the resales revenues over time. This possibility has not been exploited by any superdistribution schemes yet.

Figure 5: Examples for expected revenues from resales and effective price in random superdistribution with dynamical forward pricing.

Further results model’s analysis spark optimism for superdistribution as a business and its viability as a replacement for DRM. In a basic extension of the model it was shown in [Schmidt 08a] that the legitimate good in the CDO can prevail against a free-rider version under moderate assumptions. Nonetheless, superdistribution market mechanisms need to be carefully crafted as many more external factors other than rational decision-making based on pecuniary incentives come into play. One important aspect in that vein is market homogeneity. While superdistribution will work fine in a population which consists of a rather homogeneous group of
individuals, for instance with special preferences, it may break down if the market is biased in the sense that there is a group of agents with higher trading capacities, e.g., large music labels running direct sale web sites. Furthermore, inhomogeneities amplified by network effects [Economides 96a, b], [Swann 02], [Lim, Choi, and Park 03], [Maurer and Huberman 03], carry the imminent danger that the market can be cannibalised in the sense that a single reseller reaches a practical monopoly despite the fact that he has no differentiating value proposition. This can happen at an early stage by an agent with overwhelmingly high communication capacity, e.g., a popular web site.

Finally, there is a psychological element to superdistribution that is connected to the aleatory element of network markets and human sense of justice, which modern empirical economics has shown to be an important driving force of human action [Fehr and Gächter 00]. In the small-scale study on a real superdistribution system [Ahrens, Hess, Pfister, and Freese 08], it was shown that users felt bad about the monetary incentive they received from resales since they were asking money from their peers for something that was perceived as pure entertainment. Though these results may be culture-dependent to some extent, they show that the marketing aspects of superdistribution deserve utmost care.

4.3 The security and technical axis

From a security viewpoint the central difference between DRM and superdistribution is that DRM protection is focused entirely on the CDO, while in superdistribution the most important protection goals regard the remuneration. In fact, the parts of superdistribution which require local DRM protection in and between the nodes are encoded in the consumption rules of the consumption licence and the redistribution rules. The latter are essential to protect the business model and market mechanism implemented by the superdistribution system’s operator. These CDO protection requirements can be implemented by arbitrary DRM measures, centralised or decentralised and with a varied level of enforcement, as we have seen in the examples of Section 3. An important point for the buyer is the secure association to the content to prevent the mentioned content masquerading. On the other hand, ensuring remuneration is essential to implement a fair superdistribution market. A natural way to combine the in-band with the necessary out-of-band processes, e.g., payment, is by sending back receipts, which are cryptographically bound to the content and transaction, to the reseller. The reseller can then for instance redeem these receipts as tickets at a central rewarding service. We have shown a way to implement such general schemes with trusted platforms in [Kuntze and Schmidt 07].

Privacy is of utmost importance in a network of transactions involving a large number of partners. In superdistribution privacy is limited again essentially in the remuneration process, since there buyer and reseller need to reveal their identities and transaction data toward a payment provider or transaction processing service. This is not a gross risk to privacy, since often buyers and resellers are acquainted anyway, for instance if superdistribution is based on personal recommendation. In general, the identities of nodes in the RON should be protected by Identity Management systems [Clauß and Köhntopp 01], [Pfitzmann and Waidner 03] to the appropriate level. The Paradiso system described in [Section 3.2] exhibits the usual trade-off between security and privacy. The chain of licences transported downstream in the CDO
contains information (though not necessarily personalised) on every superdistribution transaction on a path. It would be interesting to see if security can be protected with similar strength but with higher privacy levels. Methods for that can for instance make use of cryptographic zero-knowledge proofs [Feige, Fiat, and Shamir 98], [Camenisch and Herreweghen 02].

5 Multimedia Superdistribution over Peer-to-peer-like Networks

Superdistribution offers advantages with respect to cost-efficiency, scalability, and ultimately quality and dependability of service in such systems. Content distribution solutions have evolved from classical client-server models, through distributed caching, to Content Distribution Networks (CDNs), and more recently p2p networks. Data centres, i.e., facilities which host large numbers of networked computer servers and power supplies are often critical enablers of such services. They are a major source of cost and complexity for operators, while they are inherently not scalable due to their centralised nature. As a result, router companies, server manufactures, and hosting facilities hasten to produce more efficient hardware and software for data centres, and aim to improve their efficiency of operation. Dynamical load balancing is the most widely used technique for this. While this effort improves efficiency, it is bound to produce rather short-term remedies. For these reasons, several European partners have joined forces in the EU’s seventh framework programme in the project NanoDataCenters [Nada 08]. The aim is to leverage the vast computational and storage capacities nowadays present on the edge of the network, i.e., in the users’ home in the form of, for instance residential home gateways or set-top boxes. Nada strives to combine the best of p2p-like superdistribution with advanced data warehouse technology under real-time constraints to achieve utmost efficiency and scalability under real-time constraints.

Although some commercial p2p content distribution services such as JOOST [Joost, 2007] and BabelGum [Babel, 2007] have emerged, their technological and economic viability remains to be shown. Incentives in p2p are investigated mostly on the functional level to ensure nodes’ fair behaviour. Nada requires also incentives at the user level and pertinent economic research [Zghaibeh et al. 07], [Rodriguez et al. 06], [Chen et al. 07]. Some central questions of practical relevance for large-scale, commercial p2p deployment are, however, under-researched. In particular the security and integrity of distributed content is not sufficiently covered. Finally, user privacy (personal profiles, uploaded content) requires novel concepts in superdistribution networks. Technical state-of-the-art CDN involve DRM which often do not respect privacy [Fewer et al. 07].

5.1 Security of Nano-Data-Centres

The approach of Nada naturally bears some particular security requirements. This means in particular the necessity to transcend classical end-to-end security paradigms in a step toward de-centralised security architectures. The central requirements of Nano-Data-Centres in the users’ home are these:
Security properties of nodes must be maintained although the nodes of the CDO may be off- and online in an unpredictable manner, in particular beyond a node’s boot cycle.

Nodes must be enabled to build trust relationships autonomously, i.e., without the help of a central entity.

Private data of different stakeholders in the nodes must be separated from each other and separately handled by secure functions under the ultimate control of the respective stakeholders.

Nodes must be able to autonomously build CDO and RON according to given licences and enforce the latter.

The content distribution system envisioned in Nada both involves the download of content from a centralised source or several sources to home gateways and the upload of various types of data and content from each gateway to various sinks located in the overlay network. The protection of these two types of data flows in a way that is consistent with the Nada content distribution mechanisms is a challenging problem that calls for new primitives and protocols.

One new requirement is to provide a security mechanism that allows the content distribution mechanisms to operate on encrypted or integrity protected data while preserving the security of these mechanisms and without having to trust intermediate components (nodes) involved in data distribution. The other new requirement raised by the Nada system is about the privacy of the users. As opposed to classical content distribution systems, Nada requires a strong involvement of users in basic data distribution mechanisms, ranging from providing customer-created content to uploading user preferences about distributed content. Privacy in this environment is a crucial requirement that calls for the confidentiality of the user data and the personal information of each subscriber in the face of potential disclosure by service providers, other users or intruders. The design of mechanisms that allow user-level operations while preserving the privacy of the users is thus another challenging problem.

The new data confidentiality and integrity mechanisms must be compatible with the quality of service (real-time streams) required by the content type and communication mechanisms underlying the Nada operation. Management of security-critical data and functionality from a remote needs a high enforcement level on the edge nodes and integration of hardware-based security to establish the needed trust relationships. There is only one technology in sight that offers the required guarantees on security in a de-centralised fashion. Trusted Computing (TC), [TCG 2007a, b] as standardised by the TCG relies on a hardware trust anchor and a certain set of core capabilities in every computing platform. In the following we explain the approach of Nada to security based on TC in the basic example case of set-top boxes in DVB content distribution systems. The technical realisation is described in forthcoming papers [Bal, Kuntze, Schmidt 09], [Brett, Kuntze and Schmidt 09].

5.2 Trusted Set-Top Boxes

Digital Video Broadcast (DVB), as the widest spread standard for digital content delivery, comprises already some methods for the protection of media data. DVB exists in variations for different broadcasting techniques and formats: Satellite (DVB-S), cable (DVB-C), terrestrial (DVB-T), and mobile environment (DVB-H). The signal is encrypted with the Common-Scrambling-Algorithm (CSA) using an 8 byte
seed for initialization, the so called Code Word (CW). This Code Word is provided by the Conditional Access System (CAS) [Gallery and Tomlinson 07]. There are many vendors like Cryptoworks or NDS offering CAS Modules to content providers, see [Table 7.1 in Leung, Yau, and Mitchell, 07]. The CAS has the essential task to bridge between the encrypted data stream and a smart card providing CWS. Due to various different CAS systems the customer needs different smart cards, often for exclusive use with different, proprietary Conditional Access Modules (CAM). CSA was kept as a secret over a couple of years, but was revealed some time ago [CSA 07]. Until now CSA is not broken [Weinmann and Wirt 05].

Charging and payment is another purport of the smart card – set-top box (STB) combination, a market which is dominated by smart card subscriptions. The customer registers the card after purchase with the provider and is able to descramble the digital stream for a certain time. On this basis, pay per view schemes, e.g. for single movies, can be realized. Actual charging is sometimes solved by using value added telephone services. A second way of charging for DVB content is using mobile payment solutions. One (German) peculiarity is the use of debit or credit cards in combination with a feedback channel of the set-top box [Conax 02].

The traditional DVB architectures have some common problems associated with the stand-alone nature of the set-top box, the unavailability, respectively, costliness of an upstream channel, and the smart card-based security architecture:

1. The update of decryption algorithms and secret keys and generally the remote management of the STB, e.g., subscriber management and channel bouquet selection, are difficult and costly.
2. Accounting and charging is generally realized as an out-of-band process more or less tightly linked to smart card roll out. On-line charging solutions are scarce and of provisional nature.
3. Users selecting bouquets from a variety of subscribers have to handle a number (ever increasing as the market diversifies) of smart cards manually.
4. If bouquet selection is done via the DVB down link, sending of personalized data for this kind of access control over the DVB channel is costly and does not scale well for many subscribers.

The focus here is to provide practical improvement of the existing content protection schemes used by the DVB standard providing benefits in terms of customer satisfaction and price of the individual device. Other work in this area is focused in a more general security analysis as by [Gallery and Tomlinson 05] and [Mitchell (05)]. These security-theoretic approaches may be integrated in the scenarios of this paper. Nevertheless, from an economic perspective the relevance of certain protection targets like freshness and proof of origin has to be examined with respect to the practical importance. We propose architectural concepts based on Trusted Computing (TC), to improve on this state of the matter, and present a high-level design of a trusted set-top box (TSTB) which can be reconfigured for various content protection schemes and payment methods. At the core of the concept lies a trust-enhanced CAM. In particular it realizes descrambling methods in software while protecting the associated access secrets of each provider. This usage of TC for STBs is rather traditional and feasible with minimal architectural changes. In a further step, we assume that the trusted STB has some sort of access to a communication network and use the latter for take ownership of it, i.e., the process of impressing a user identity and associated
credential to the STB. This essentially obliterates the use of smart cards. Finally we discuss options for integration of charging functionality using a mobile device communicating with the STB.

5.3 Trust-enhanced CAM

Building on the basic features of TC a soft realization (or even virtualization) of the CAS is feasible. Implementing in this way a TSTB can be implemented in many variations. An elementary implementation stores the functionality of the CAM as software protected by means of the Trusted Platform Module (TPM). If the user requests a scrambled channel, the TSTB uses the CAM software to create the respective CWs required by the CAS. A system architecture is shown in [Fig. 6].

Starting with the smart card on the right hand side, we observe a main conceptual change w.r.t. traditional CAS: The smart card bears merely the subscriber identity and credential. All other security-critical functionality is shifted into the trust boundary of a trust-enhanced CAM and protected by a hardware security anchor, e.g. a TPM. The TCB enables and protects upper functionality of the CAM via a secure boot process. The protected CAM functionality comprises three essentials: i) A management software and protected, non-volatile storage for access secrets used by the single CW generators. ii) An interface to the smart card providing a secure channel to import user credentials iii) A secure interface to import algorithms and access secrets.

The access credential manager invokes soft instances of CW generators on demand, using the required access secret, e.g., a Control Word. Security of the CAS realized in this way relies on the fact that CW generation is within the trust boundary,
i.e., the system state of the CW generator and access secret manager is trustworthy and tamper-resistant.

The TCB of each Set-top box is equipped with a credential which identifies it as belonging to some group or individually. This enables end-to-end (e2e) encrypted communication from the provider’s headend [ATIS 07] to the Trust-enhanced CAM. This credential can be impressed at an early stage of the STB’s lifecycle, e.g., by the manufacturer, the OEM, or a service provider. It can be located in the hardware trust anchor or be protected by it.

New access secrets and usage policies can be transferred to the CAM in the common Entitlement Control Messages (ECMs) and Entitlement Management Messages (EMM) respectively [3GPP 03], distilled from the DVB stream [ETSI 96], [ETSI 05]. They are, in our scheme, e2e-encrypted for the particular CAM and enable the management software to fell authorization decisions. Comparing the subscriber identity with the policies for e.g., bouquet access in the EMMs, the manager can decide whether to import the associated Control Words which embody the access secrets for the bouquet. Algorithms can be updated for each CW generator, e.g. via IP-over-DVB [ETSI 03].

Some key advantages of a trust-enhanced CAM are obvious. It resolves the focal problems 1. and 2. described above. The CW generators can be realized in software and implement a variety of different CW generation algorithms. These algorithms can be updated “over the air”. Additionally it has the technical advantage that the frequent sending of ECMs for security reasons becomes obsolete, since the derived access control secrets, commonly today the Code Words used for initialization of the CSA, are managed inside the trust boundary and hence are not easily accessible to attackers. This helps to address problem 4.

Let us discuss algorithm updates in some more detail. CSA is designed to meet two central technical requirements. First it has to provide a secure scrambling of the data. Second it has to operate in a real-time environment. The digital stream has to be descrambled the moment it arrives. This leads to a lack of algorithm strength and the need to replace the algorithm can be foreseen [Weinmann and Wirth 05], [Bewick 98], entailing a complete change of the installed infrastructure on side of the customers as well – a very costly and inconvenient effort.

The trust-enhanced CAM enables a modular system in which it is possible to replace every part of the descrambling mechanism in a trustworthy way. Using for instance Field Programmable Gate Arrays (FPGA) technology on top of TCB and root of trust enables the replacement of every part of the existing DVB architecture. For instance, algorithms can be implemented in hardware to gain speed. FPGAs are programmed (reconfigured) before they can perform the desired task. Trustworthy implementations can verify the content of the FPGA before data are transmitted between FPGA and the CAM.

In a more evolved scenario the root of trust, e.g. a TPM, could be an integral element of the code word scheme replacing the smart card.

### 5.4 Online CAS

A system which emulates the actual CAS systems in software has to enforce that keys used to generate the CWs required by the CAS are kept secret. The security of the proposed system has to be guaranteed even if the algorithm is published. Beside this
technical requirement online verification of the access rights is a premise for
dynamical content access control. This leads to an authorization scheme where the
CAS asks for permission before the CWs are created. Therefore a connection is
established by a communication module granting access to a network.

Figure 7: Online CAS Scenario

A protocol solving this problem has to perform the following steps. In preparation
an appropriate, controlled roll out of the STB is required. During the roll-out a take
ownership by the provider must take place. A network operator can be used instead of
the provider assigning user identities and, optionally, performing charging processes.
The proposed scenario consists of four parties: customer, provider, charging provider,
and network operator as depicted in [Fig. 7]. The network operator issues the set-top
boxes to the customers in the same way as they offer mobile devices. The bonding
between customer and device is based e.g. on a SIM-card. An online take ownership
is as well possible as described in the subsequent section.

The network operator establishes his trust in the device by a network operator
issued credential which is produced by the box on request. Based on this underlying
trust relation the (M)NO can assure the identity of the set-top box to a supplier. This
second (transitive) relation [Kuntze and Schmidt 06] is based on a second credential
issued either by the network operator or the trusted set-top box. In either case the
network operator in his role as an identity provider (ID) signs this credential which
therefore is stored in the trusted set-top box. If a customer decides to consume a
certain service it offers this credential to the vendor (V), for instance the content
provider. V uses this credential to verify the identity against ID. V then delivers the
content and requests charging by ID. In this scenario the user is unknown to V as the
credential is only validated by ID. ID does not need to reveal the user identity to V.

Delivery in this context means that V transfers a secret to the TSTB and adds this
secret to the list of accepted credentials as this is known by the actual process of
conventional CASs. This secret is sealed in the set-top box by using the sealing
functionality of the TPM. This means that it can only be used in the same trustworthy state of the box. The root of trust in this case is the ability of attestation of the integrity to a third party namely V and/or ID. An advantage is that personalized data, e.g., for bouquet access needs not to be transferred via the costly satellite downlink.

5.5 **Online Take Ownership / Online Registration**

The aim of an online take ownership procedure is to establish a user identifying credential in a TSTB without the need to issue the box over a special infrastructure provided by ID, in contrast to the previous subsection. The customer should be able to buy such a box everywhere he/she wants. During production every box gets an identifier in form of the unique platform certificate. Based on this initial credential a protocol can be performed to establish a user related certificate. This user certificate identifies the user at V. The used protocol establishing this user credential depends on the existence of a direct or only indirect communication between V and ID.

The user certificate can be created after the take ownership process of a trusted platform which binds a TPM to a certain user using a 160 bit authentication value (TPM owner authorization). So called Attestation Identity Keys (AIKs) are available after the take ownership to be used as credentials testifying the trustworthiness of the state of a platform without revealing the identity of the platform or its user. An AIK can only be created offering a valid TPM owner authorization and is a private/public key pair. The private portion is shielded inside the TPM. After this, a Privacy CA (PCA) issues a certificate to assert the security association between AIK and TPM. For this AIK and certificate creation process an online connection to a privacy CA is required. The pertinent protocol has to protect the origin of the key so that it is impossible to fake a TPM.

![Figure 8: Online take ownership via mobile device and network.](image-url)

In the case of a direct communication the system is equipped with a communication device enabling the direct contact to the PCA. In this case the mentioned protocols can be used without any restrictions. If the system is not equipped with such a communication device at least a short range communication is required enabling a take ownership over a secondary communication device (SD) like
a cell phone. The SD forwards the communication in the respective direction. The ensuing communication relationships are shown in [Fig. 8].

It is important to mention that there has to be a trust relation between the PCA and the content providers. In this use case AIKs are used as tickets which enable accounting. Therefore it could be possible that the PCA should be able to reveal the identity of a certain AIK, e.g. in case of suspected fraud.

After the take ownership an online registration at the respective provider is required to sign up to a certain service. In this process two goals have to be achieved. First, the identity of the mobile device (and therefore the identity of the user) has to be registered at the service provider. By issuing the AIK and the associated certificate, the identity can be proved, and by performing a handshake protocol between service provider and mobile device the origin is testified. The second aim is to negotiate the conditions of the subscription, of which the payment information is the most important part. The service provider transmits a data structure which describes the available charging models and services. The user selects from this offer, signs the selection with the private portion of the AIK and transmits this to the service provider.

A proof of authenticity of the service provider is also required. Hence it is necessary to sign messages issued by the service provider. A verification of the authenticity can be achieved using known PKI structures or built-in root certificates. Alternatively, for instance a Mobile Network Operator (MNO) can vouch for the authenticity of a certain service provider replacing conventional PKI systems.

![Figure 9: AIK authentication for content access.](image)

[Fig. 9] shows a basic scheme using AIKs as authentication tokens. The protocol is divided in two phases. Phase 1 is concerned with AIK creation and certification, phase 2 is the usage of this token. The set-top box transmits in 1) certain credentials of the platform and the public portion of the AIK to the PCA which verifies the offered credentials and then retransmits (2) a certificate stating that this AIK belongs to a an accepted platform. If the MNO works as PCA the AIK can later be used as a payment credential as the MNO can reveal the identity of the users based on the certificate. This feature is used in phase 2 where the set-top box offers the AIK and the corresponding certificate to a provider (3). This provider performs an attestation of the box in this step and then requests from the CA the validity of the offered certificate (4,5). 4 and 5 can be performed e.g. using OCSP responder. Step 6 returns the signed acceptance information of the provider. To validate the signature, an appropriate certificate must be available to the TSTB. It can be necessary to request this root certificate from a trustworthy third party. After this the TSTB has been
6 Conclusions

The main claim of the present paper is that superdistribution is conceptually different from both DRM and p2p and is a third field in its own right. In fact we have shown that the system theory of superdistribution is much richer than for DRM systems. Superdistribution uses — by necessity — informational representations for the value proposition of a digital good to its buyers, i.e., the combination of consumption and remuneration for resales. Moreover the economy of superdistribution lies on a categorically different level than the economy of p2p networks, which is centred on questions of incentives for participation and fairness in the contribution of resources [Antoniadis, Courcoubetis, and Mason 04], rather than transported values.

This has important socio-economic implications, since superdistribution allows for business models for that can live without restrictive DRM measures. The conflict between copyright holders and free-riders or “pirates” has led to distortions in economy and legal regulations which diminish consumer experience and arguably even inflict on personal freedom. Successful superdistribution models could be a bridging element to restore a balance of interests in this context. A full discussion of this is contained in the workshop version of this article [Schmidt 08b].

Superdistribution as such is almost technology-neutral. Three challenges need to be met for their success in the economy of digital goods:

- **Market mechanisms** must be implementable in a general superdistribution framework or platform. Such a framework should enable the definition of CDO and RON, for instance rewarding levels, match-making rules, allowed number of resales, or the more concrete rules some of which have been mentioned in [Section 3.3].

- A **marketing platform** must be incorporated in the network, in particular to ensure fairness in trade and competition between resellers, and market homogeneity.

- The **dynamisation** of the market should be supported. This regards local changes in space and/or time of the two licences, of which perhaps the most important example is dynamical forward pricing. A related research challenge is to devise methods to monitor the market in real time. This would for instance be useful to furnish up-to-date information on the popularity of a piece of content.

- As an example, the digital good could be made returnable to the originator or the reseller if the chances to achieve further resales revenues becomes too low.

We conclude that the evolution of superdistribution based business models for digital goods is still in its early beginnings — and though the risks are considerable, the prospects are equally thrilling. As a research subject, superdistribution can be really attractive since it is interdisciplinary by nature and at the same time has a clearly defined field of experiment in the digital economy.

Acknowledgements

Special thanks go to Shiguo Lian for inviting this contribution to the MUSIC’08 workshop at CHINACOM. I thank the partners in the NanoDataCenters EU project and in particular Nicolai Kuntze for their collaboration.
References

[2peer 08] 2peer: https://www.2peer.com/ Visited 07.12.08


[4FriendsOnly 08] 4FriendsOnly AG: http://www.4fo.de/ Visited 20.12.08


[EFF Grokster 08] Electronic Frontier Foundation page on Grokster: http://w2.eff.org/IP/P2P/MGM_v_Grokster/


[ETSI 05]. ETSI TS 102 367 V1.1.1: “Digital Audio Broadcasting (DAB); Conditional access.” 2005


[IDMT 08] Fraunhofer IDMT: http://www.idmt.fraunhofer.de/ Visited 09.12.08


[Potato 08] Potato system: http://www.potatosystem.com/ Visited 07.12.08


[SpreadBox 08] SpreadBox: http://www.spreadbox.net/ Visited 13.12.08


