The Threat of Social Media Blackmailing in the Hospitality Industry – when Customers misuse their Power

Author: Kelly M. R. Raas
University of Twente
P.O. Box 217, 7500AE Enschede
The Netherlands

ABSTRACT
Growing popularity of social media has lead to important challenges for reputation management given the interconnected online environment, where the empowered consumer is taking the lead. Nowadays customer ratings are among the most decision-relevant information on the Internet. Reviews are particularly valued for their credibility compared to mass media advertising and are increasingly used by customers to make purchase decision. Due to the high relevance of the review sites there is a growing concern for the repercussion negative reviews can have on a businesses reputation. The high relevance of customer reviews on market leaders such as TripAdvisor or HolidayCheck is especially of concern in the hospitality industry. In the past years a growing number of cases where reported where clients are trying to extort hotels by threatening with a bad review, unless demands like refunds and upgrades are fulfilled, this phenomenon has become known as social media blackmailing.
This research was performed in order to set this new arising type of online reputation threat in the context of online reputation attacks, in order to provide a clearer picture about the kind of attacker who engages in social media blackmailing and its underlying motivations. By conducting a critical literature review of scientific theories as well as discussing and analysing articles of non-scientific nature, a solution, given as a four stage action model, for countering blackmailing threats is provided. The mode is further supported by findings from two cases.

Supervisors: Dr. E. Constantinides, Dr. R. P.A. Loohuis

Keywords
Online reputation management, online reputation threats, hospitality industry, review systems, blackmailing, false online reviews
1. INTRODUCTION

1.1 Background: Corporate Reputation

In today's competitive markets, companies realize that the role of reputation is becoming increasingly important. As Walker (2010) points out, the relationship between reputation and sustained competitive advantage is widely acknowledged in the literature (Fombrun and Shanley, 1990; Fombrun, 1996; Hall, 1993; Roberts & Dowling, 2002). Some scholars, such as Gibson et al. (2006) even suggest that corporate reputation is the single most important organizational asset. A good corporate reputation can lead to a number of strategic benefits, such as lower firms costs, ability to charge premium prices, attracting applicants, investors and customers, higher profitability and creation of competitive barriers (Walker, 2010). Consequently, it is focal for a company to build and sustain a good corporate reputation.

However, today in the era of social media, companies have ever less control over their reputation. It is no longer defined by what they “report” they stand for, but instead it is increasingly defined by the shared opinions and experiences of the socially-connected consumers (Bulmer & DiMauro, 2014). With consumers gaining increased influence on corporate reputation, threats for reputation management reaches a new degree of severity, as some consumers take advantage of their new power.

The misuse of customer power is the focus of this research, therefore the next section provides further insight into the context of the subject and introduces the research problem.

1.2 Research Problem: Reputation Threats

"Give me what I want, or I’ll give you a bad review or write a negative blog post about you." A sentence growing number of business owners fear for today as social media or review blackmailing is becoming the new threat in the world of social media.

According to a recent article in the British newspaper “The Telegraph” ("hostage to TripAdvisor blackmailers", 2014), more and more hotels and hospitality executives are being targeted by customers who demand upgrades, freebies or other special treatments in exchange for not writing a negative review. In fact, these customers are blackmailing the business owners. Correspondingly, a similar report was published in the Dutch "De Telegraaf" ("Afgesperd op Internet", 2014) where CEOs within the Dutch hotel and restaurant sector discuss the same issue of costumers frequently blackmailing businesses for their own advantage.

Today, in the era of Web 2.0, social media provide an unequalled platform for consumers to disseminate their opinions about products and services and making it accessible to thousands of other peers — i.e. through word-of-mouth or consumer reviews. Moreover, there is increasing evidence that online word-of-mouth has a significant influence on purchase behaviour (De Broun & Lilien, 2008; O’Connor, 2010). For example, a survey by Opinion Research Corporation (2008) found that over 60% of respondents checked online reviews, blogs, and other customer feedback before buying a new product or service, and over 80% of those who consulted such material said that it had at least some influence on their purchase decision (O’Connor, 2010). Hence, the threat of a negative review in today’s environment is quite significant.

Before the social media gained the importance and influence it currently holds, consumers were limited to the recommendations of agents, friends or families based on their individual perspectives. Now with the web 2.0 the concept of word-of-mouth has reached the dimension of a mass communication medium. And with the opportunity to spread word-of-mouth globally, consumers gain increasing power to impact the success and reputation of a brand or firm.

Mazzarol, Sweeney, and Soutar (2007) postulate that such electronic word-of-mouth (eWOM) is thought to be particularly important for services, as their intangible nature makes prepurchase trial impossible.

As such, the hotel industry is particularly vulnerable to review threats, as an increasing number of bookings are done online, the decision process is likely to be influenced by what other previous guests post on one of the many travel review sites such as tripadviser.com or holidaycheck.com (Sparks & Browning, 2010)

Thus, it is no surprise that hotels and other hospitality businesses are increasingly worried about customers holding them to ransom with the threat of a bad review unless they get a refund, an upgrade or a free meal. The available evidence throughout the media suggests that this is widespread and growing problem.

However, scientific literature on this particular topic is still scarce. Despite plenty of research on reputation management and how to cope with negative customer reviews there is a gap when it comes to reputation threats such as social media blackmailing. Information on how often social media blackmailing occurs or how often consumers make their threats become reality would be valuable in order to develop strategies to either prevent or deal with social media blackmailing.

Therefore, the aim of this paper is to assess the literature on the issue of social media blackmailing in order to gain a deeper understanding of the matter and provide a systematic approach for businesses to deal with the problem, and finally bringing this topic to the attention of researchers by providing ground for further research.

In order to do that the following research question will be answered in this paper:

“How can businesses deal with blackmailing threats in the social media environment”?

The relevance of this study lies in the fact that there is important support for the claim that customer power and influence in the social media will keep growing. Social media blackmailing therefore represents an important and growing issue, threatening the reputation of businesses. Given the centrality of this issue it seems necessary to provide more insight into this topic through research in order to come up with strategies for businesses to protect themselves. This paper aims to provide a basis for an academic contribution to the literature, as well as a practical contribution for businesses.

In order to give an adequate answer to the research question, the following sub-questions are discussed:

1. What impact do online reputation threats have on companies?
2. How do reputation threats differentiate - in terms of motive and style?
3. What strategies can companies use to deal with social media blackmailing?
1.3 Methodology
A critical literature review is conducted in order to identify the current state of knowledge on the issue of social media blackmailing and to provide a general insight into the topic. Social media blackmailing per se is a quite recent phenomenon, therefore scientific research on this matter is quite scarce. However, it is part of the broader subject of online reputation threats, thus literature of this field is used to provide an underlying understanding of the issue. Additionally, in order to develop a ground and direction for further study on the subject of social media blackmailing, non-scientific literature, such as newspaper articles, blog entries and other online media is scanned to provide the most actual information about the phenomenon.

Moreover, a case study is presented to validate the theoretical findings from the literature review with practical findings from the hotel industry. As a form of research, the case study is highly valued for its ability to consider a single or complex research question within an environment with many contextual variables and is best suited to examine the how and why questions (Schell, 1992). Thus, it seem the most suitable approach when considering the questions why social media blackmailing occurs and how businesses can deal with it. For this matter, two interviews are conducted within the hotel sector. The case studies are analyzed in order to highlight the strategic approaches on how the business deal with blackmailing threats. The practical approach from the cases are then compared with the theoretical findings from the literature review.

1.4 Research Structure
This paper considers the phenomenon of social media blackmailing by investigating the main sources and elements of online reputation threats in general.

After an introduction to the key concepts, the paper is structured as follows. It introduces the challenges of online reputation management (Section 2.1) and provides a review on motives and styles of online reputation attacks (Section 2.2). Following, this knowledge is used to reflect on the phenomenon of social media blackmailing. After that, some strategic approaches to deal with the issue are introduced (Section 2.3). To validate the empirical findings, two case studies drawn from the hotel industry are analyzed (Section 3). Finally, the findings are discussed (Section 4) and recommendations for further researched are proposed (Section 5).

1.5 Definition of Key Concepts
In order to facilitate the understanding and provide a good lead-in to the subject of interest in this paper, first the basic key concepts are defined.

1.5.1 Social media
To provide a comprehensive concept of what is meant by the term social media, it is necessary to point out the relation it holds with two related concepts that are often used in conjunction with it, namely Web 2.0 and User Generated Content.

As Kaplan and Haenlein (2010) explain, the term Web 2.0 was first introduced in 2004 to describe a new way software developers and end-users started to use the Internet. Instead of merely being a platform where content and applications is created and published by individuals, it turned to a highly interactive hub where content and applications are continuously modified by all users in a participatory and collaborative manner. Though, Web 2.0 does not refer to a specific update of the Internet, there are a few technical features necessary for its functioning, such as Adobe Flash, RSS (Really Simple Syndication) or AJAX (Asynchronous Java Script). Web 2.0 can therefore be considered as the technological foundation for the development of Social Media (Kaplan & Haenlein, 2010).

User Generated Contend (UGC) is a fundamental element of social media. The term achieved broad popularity in 2005 and is usually applied to describe the diverse forms of media content created and shared by end-users (Kaplan & Haenlein, 2010). With UGC users add value to web applications and processes by generating, reviewing, editing and disseminating content (Constantinides, Romero, & Boria, 2008).

Based on these clarifications, social media can be defined as the new generation of internet based tools and applications, such as blogs, wikis, online communities, and virtual worlds, where users are in control of the content, and used to build social connections, collaboration, affinity, and relationships (Constantinides et al., 2008; Parise & Guinan, 2008).

1.5.2 Consumer empowerment
The concept of consumer empowerment has been widely discussed, with some variations in its definition depending on the different perspective used.

The current conceptualization of consumer empowerment becomes a threat for corporate reputation. With the control of market information shifting towards the consumer, businesses have less control, if any, about the information that is being created about them and customers have increased ability to harm a firms reputation.

1.5.3 Corporate reputation
Despite the importance of the concept of corporate reputation, the definition of the term still is a fundamental problem in academic literature. Yet, different disciplines prefer different terms, for example sociologists use the term prestige, economist prefer the term reputation, marketers favour “image” and accountants talk about corporate goodwill (Wartick, 2002). In particular the terms identity, image and reputation are often used interchangeably (Barnett, Jermier, & Lafferty, 2006). At least most researchers seem to agree on the major differentiation between the three terms, which is whether the concept refers to the perception of internal or external stakeholders, or both. While corporate identity refers to internal stakeholders and corporate image to external stakeholders, corporate reputation refers to both internal and external stakeholders (Walker, 2010). Another distinction pointed out by Walker (2010) is that while corporate identity and corporate reputation are actual perceptions and therefore could be positive...
as well as negative, corporate image is a desired perception and therefore always positive. With the general understanding of how corporate reputation differs from identity and image, an overall definition of the term can be approached. In simple terms, Balmer and Greyser (2006) describe corporate reputation as the answer to the question “What are we seen to be?”. Drawing on this question, for the sake of this research, we define corporate reputation as “the collective of internal and external actual perceptions of a companies’ actions and behavior.”

1.5.4 Reputation System
With the Internet consumers have gained a nearly boundless access to information. However, increased quantities of information can be both a blessing and a curse (O’Connor, 2010). The sheer quantity of information available can be overwhelming and complicate the decision-making process, when the abundance of alternatives leads to confusion or dissatisfaction with the choices made.

To overcome these challenges, one of the most ancient mechanisms of human society, word-of-mouth, is gaining new significance in the cyber space, leading to the emergence of reputation systems. These are web-based opinion platforms that collect, distribute, and aggregate feedback about anything from products (Amazon product rating system), businesses (TripAdvisor hotel rating system), users (sellers and buyers on eBay) or digital content (video clips on YouTube).

1.5.5 Social Media Blackmailing
Problems emerge when customers misuse their social media power in ways threatening the reputation of businesses. This can happen in the form of various online reputation attacks, which have become some of the easiest and most impactful mechanisms used by individuals to cause serious harm to businesses.

One new breed of such attacks is social media blackmailing, especially in the service sector such as the hotel and hospitality industry. There is no official agreed definition for the term of social media blackmailing. Hence, for the purpose of this article, the definition provided by TripAdvisor for ‘review blackmail’ is adopted. It states “when a guest threatens the management to post a negative review unless demands like refunds and upgrades are fulfilled”.

2. LITERATURE REVIEW

2.1 Social Media and the Challenges for Reputation Management

Reputation, as an important source for sustained competitive advantage, is a critical element in strategic management. However, the new Web 2.0 environment challenges reputation management as social media expands the spectrum of reputation risks and boosts risk dynamics (Aula, 2010). As Aula (2010) points out, in the social media environment, content can neither be controlled nor managed in the same way as traditional media such as TV or newspapers. In practice, this means that it is almost impossible for organizations to control the information spread about themselves.

As Bulmer and DiMauro (2014) put forward, the reputation of a company is no longer defined by what they “report” they stand for, but instead it is increasingly defined by the shared opinions and experiences of the socially-connected consumers. Hence, it can be argued that today a company’s reputation is its online reputation (Fertik & Thompson, 2010)

While companies have to invest a lot of time and money to build their reputation online, the same companies can see their reputation damaged or destroyed in a short period of time. Due to the nature of social media, crises can disseminate rapidly and create long-term damage to a company’s reputation in a matter of minutes (McCorkindale & DiSato, 2013). As a famous quote from Warren Buffet goes “It takes 20 years to build a reputation and 5 minutes to ruin it”. According to Walsh (2002), “reputations are made or destroyed online by any of the many different kinds of stakeholders who now have inexpensive and effective means to do it” (p. 38). Virtually anyone with a motive can potentially harm a business and its online reputation by publishing a series of damaging false reviews. In a time where consumers increasingly rely on social media and review sites for decision making, this is an important issue.

As apparent, social media poses a constant threat to a company’s reputation if it is not managed properly (McCorkindale & DiSato, 2013). Scholars agree that it is critical for companies to invest in proactive communication and appropriate response strategies to protect their valuable yet fragile corporate reputation (Aula, 2010; Fertik and Thompson, 2010; McCorkindale & DiSato, 2013). For example, by developing strategies and guidelines for monitoring, understanding and responding to different social media activities (Kietzmann, et al., 2011).

2.1.1 The Power of Online Reviews

Online reviews have gained increased significance and can have a great impact on a company’s reputation. Interpersonal influence and word-of-mouth are ranked the most important information source when a consumer is making a purchase decision (Litvin, Goldsmith, & Pan, 2008). As O’Connor (2010) points out, in an information intensive situation, consumers actively seek out the opinions of others as a way to reduce uncertainty and risk.

The advantages of the internet and proliferation of social media have profoundly changed the way information is transmitted and have transcended the traditional limitations of WOM, leading to the emergence of the term eWOM. This concept of eWOM is defined by Hennig-Thurau, Gwinner, Walsh, and Gremler (2004) as “all informal communications directed at consumers through Internet-based technology related to the usage or characteristics of particular goods and services” (p.39). Key attributes of eWOM are that it allows individuals to share their personal thoughts, reactions and opinions in a global manner (Dellarocas, 2013) that remains available over a long period of time, and tends to be anonymous (Hennig-Thurau et al., 2004).

In this regard, consumer reviews and ratings are the most accessible and prevalent form of eWOM (Chatterjee, 2001) They are perceived as particularly influential as consumer seem to prefer such peer recommendations over other forms of input (Senechal & Nantel, 2004). This can be explained by the fact that those opinions, written from a customer’s perspective, provide an indirect experience for the reader which is found to be more relevant and unbiased, and therefore more credible than advertisement or information provided by professionals or marketers (Smith, Menon, & Sivakumar, 2005).

Further, evidence supporting the notions of the increasing role of consumer reviews in consumer decision making is given by different studies. For example, Gretzel, Hyan-Yoo, and Purifoy (2007) quote statistics from a Forrester study that suggests that
more than 80% of online consumers say they use other consumers' reviews when making purchasing decisions. Similarly, they report statistics from eMarketer suggesting that almost six out of ten consumers prefer websites with peer-written reviews and that websites with online reviews show higher conversion rates. This also supports the claim that peer recommendations are preferred over marketers' recommendations.

Combining the previously mentioned about the consumer power given by the use of social media, in conjunction with the influence user generated contend, particularly eWOM, can have on consumers purchase decision, it becomes obvious how online reviews can have a substantial impact on businesses.

Now what makes especially hotels and restaurants so vulnerable to online reputation threats?

Electronic word-of-mouth is thought to be particularly important for high involvement products or services, yet given their intangible nature they can not be evaluated before consumption (Lewis & Chambers, 2000; Mazzarol, Sweeney, & Soutar, 2007).

Travel is such a high involvement product, in addition the hospitality and tourism industry is intensely competitive (Litvin et al., 2008) and an increasing number of bookings are nowadays done online. Therefore, it can be expected that online consumer reviews are having an important role in online travellers’ decision making (Sparks & Browning, 2010).

Statistics support this claim, as a survey by TripAdvisor reveals 89% of global travellers consider reviews as influential when choosing where to book and 96% of hoteliers believe that ratings are influential when generating bookings (TripAdvisor, 2014). Like other forms of eWOM, travellers usually perceive reviews as being more likely to provide up-to-date, enjoyable and reliable information (Gretzel, 2008).

Given that the industry is relying more heavily than ever on online reviews hoteliers are aware of the importance of blogs and reviews for their business. Based on consumer reviews and posts, a property can either gain popularity or it can lose its reputation. Accordingly, hotels and other hospitality businesses are increasingly worried about negative online reviews, and often out of the fear for a bad online reputation they see themselves forced to give in to whatever sleazy customer demand. Therefore, it is no surprise that particularly this sector is seemingly attractive for those greedy consumers, aware of the business owners fear for bad online reputation, and eager to misuse their social media power for personal gains.

### 2.2 Online Reputation Attacks

The Web 2.0 and its communication capabilities has empowered consumers to voice their complaints with ease and to share their negative experience with a multitude of other consumers (Hong & Lee, 2005). And due to the anonymity of the Internet and low social, they often do so in a unhesitating manner (Gelb & Sundaram, 2002). With just one click, consumers can post their complaints in the form of negative eWOM on the Internet (Noort & Willemsen, 2011).

It is important to distinguish that not all negative eWOM is necessarily a customer attack. However, with the great number of consumers that can be reached with negative eWOM, online complaining or negative reviews can easily be misused as a deliberate action to attack and harm the reputation of focal companies (Hennig-Thurau et al., 2004; Hong and Lee 2005; Ward and Ostrom 2006). Such malicious customer attacks create some of the most severe, pernicious, and persistent forms of reputation damage (Fertik & Thompson, 2010).

#### 2.2.1 Styles and types of online reputation attacks

It is crucially important to react to reputation attacks in order to try to limit or mitigate the effects of negative content published online. In fact, before starting to react, companies should consider the different types and styles of attacks, as different response strategies are needed. Therefore, this section reviews the styles and types of online reputation attacks, starting with the underlying motives.

Findings from research suggest different motivations for customers engaging in either positive or negative word-of-mouth (WOM). The motivations suggested by the research of Sundaram, Mitra & Webster (1998) give a clear overview. Although their study was limited to products and the traditional concept of WOM, more recent research provides ample support for their categorization also in the area of eWOM (Walsh et al., 2004; Hennig-Thurau et al., 2004; Sparks & Browning, 2010).

Whereas for positive WOM motivations such as altruism (helping other customers in making a good purchase decision), product involvement (vent positive feelings about a purchase experience), self-enhancement (seeking appreciation) and helping the company (possibility to improve) are identified (Walsh et al., 2004; Hennig-Thurau et al., 2004; Sparks & Browning, 2010; Sundaram et al., 1998). When it comes to negative WOM, most commonly identified motivations are altruism (prevent others from having a bad experience), vengeance (punishment for the negative experience) and anxiety reduction (vent dissatisfaction) (Hennig-Thurau et al., 2004; Sparks & Browning, 2010; Sundaram et al., 1998).

Based on these motivations several attempts to classify different types of “attackers” - these are people who deliberately spread negative WOM - have been made. For example, in their book Attack of the customers: Why critics assault brands online and how to avoid becoming a victim, Paul Gillin and Greg Gianfort (2012) speak of four types of “attackers”, (1) the casual complainer, who simply complains when having a bad experience but does not engage in dramatic actions, (2) the committed crusader, who is described as the motivated and successful attacker, motivated by environment, human or animal rights, (3) the indignant influencer, who is very powerful and has a high ability to influence others, and (4) the extortionist, who is motivated by personal gain like economic rewards, damage or compensation. On similar lines, a recent article by Wysocki, Kepner and Glasser (2008) describes five types of complainers, (1) the meek customer, who generally does not complain, (2) the aggressive customer, who readily complained, often loudly and at length, (3) the high-roller customer who has high expectations and usually complains in a reasonable manner, (4) the opportunist costumer, whose goal is not to get the complaint satisfied but, rather, to win by getting something the customer is not entitled to receive and (5) the chronic complainer customer, who is someone who is never satisfied and always finds something to whine about.

Gibson (2014) explains that when speaking about malicious reputation attacks, most common “attackers” are competitors, dissatisfied employees, unhappy customers, discontented investors or other people or businesses who are upset with a company and want to cause that company serious damage. Accordingly, revenge and venting dissatisfaction appear to be the dominant motives with the ultimate goal to harm the company.
The style of these attacks can vary, thus attacks appear in many forms, such as defamatory posts on complain websites, false information on social media, false reviews, hate websites, etc. There is plenty of research on how businesses can manage reputation attacks in a preventive and responsive manner by having the right social media strategy in place, this is not part of the scope of this research, for instance reference can be found in the works of Paris and Guinan (2008), Kietzmann (2011) as well as McCorkindale and DiSato (2013).

2.2.2 Social Media Blackmailing: a special case
When it comes to reputation threats such as social media blackmailing, considered a new emerging type of online reputation attacks, the attackers are less likely to be motivated by the aforementioned, but rather engage in negative WOM for the merely purpose of personal gains, using it as a form of reputation threat.

Although, the research on this type of online reputation attacks is still scarce. Some authors already identified this type of self-interest motivated attacker. As mentioned before, Gillin and Gianforti (2012) name the “extortionist” as one type of attacker motivated by personal gain like economic rewards or compensation. Similarly, Wysocki et al. (2001) speak of the “opportunist costumer”, described as a customer whose goal is not to get the complaint satisfied but, rather, to win by getting something the customer is not entitled to receive. Craig (2013) describes this type of attacker as a new emerging, challenging type of customer who, overwhelmed by its social media power and all too aware of the businesses vulnerability when it comes to online reputation, requests or outright demands concessions or special treatments with the underlying threat of posting a negative online review.

Whitfield (2014) identifies two ways these attacker engage in social media blackmailing. On one side there are customers who are genuinely visiting businesses and then threaten to leave a false account of their experience, on the other side there are people who are leaving speculative negative reviews in the hope that they will be offered a complimentary incentive to remove their feedback. This second type of blackmailing threat can take serious forms as explained by Ferti & Thompson (2010), some particularly unscrupulous businesspeople, who discovered this lucrative form of pressure, set up so called shame review sites that are designed to encourage users to post harsh critiques of businesses, make sure the negative reviews appear in search engines, and then offer the business under attack an expensive “investigative service” to remove the negative review.

As discussed earlier, especially the hotel and hospitality industry are increasingly targeted by this type of attackers, given their high online exposure and dependence on online reviews. Therefore, the question of what these businesses can do to protect and defend themselves against these kind of attacks is an urgent matter.

Academic research on this topic is still scarce, the attempt of the next section is to provide an overview of mechanisms and procedures on how to handle this threat.

2.3 Strategic Approaches to Social Media Blackmailing

The strategic approach to handle the issue of social media blackmailing is considered from two perspectives, first taking into account the reputation system, which is the tool for the threat, analysing to what extend there is the possibility of detecting false reviews once posted on the reputation system. Second, looking at the business perspective, analysing strategies of businesses in the hospitality industry that can be applied to deal with such threat.

2.3.1 Reputation System Perspective
There are basically two forms of reviews: the rating, often in form of a scale, and the written review. It is important to point out that the detection of false reviews in general is a very challenging task. However, ratings, which are numerical, are easier to handle than written reviews. According to Hines (2007) there are even online instructions on how to write a false review that looks real and convincing.

There are three dominant approaches that reputation systems can apply to detect and handle false reviews: (1) increasing the cost of false reviews, (2) detection of false reviews, and (3) mitigating the effect of false reviews.

1. Increasing the cost of false reviews
The cost for posting the review in a given reputation system is given by whether users are required to provide certain credential of having purchased a product before reviewing it (high cost) or not (low cost). Some sites, such as TripAdvisor, allow anyone to post reviews about any hotel, making the cost of review basically zero. Whereas other sites, such as Zoover.nl require the purchase of the hotel room before a review can be written, which consequently raises the cost of review. The study conducted by Ott, Cardie and Hancock (2012) suggest that false reviews (or deception in their words) is most prevalent in communities with a low review cost. Increasing the cost of reviews, by for example requiring a booking confirmation, assures that only customers who genuinely stayed at a hotel or restaurant would have the opportunity to provide feedback, which according to Whitfield (2014) would eliminate the threat of speculative attackers.

However, apparently not all reputation systems favour this approach. TripAdvisor, the most influential reputation system in the hospitality industry, for example, states clearly that they do not support this as it would stand against their philosophy that every customer should have a voice and not only the ones who pay the bill, therefore they support a self-directed review system (TripAdvisor). And on the other side, those sites that do guarantee the authenticity by requiring a booking confirmation, such as Expedia, often have too small of a sampling of guests to compete with sites such as TripAdvisor.

2. Detection of false reviews
Mechanisms to detect false reviews are often systems that study statistic features of feedbacks. Most of them detect false reviews based on the majority rule, which considers the ratings far away from majority’s opinions as false reviews. For example, the extended Beta-function model, originally proposed by Josang and Ismail (2002) and extended by Whitty et al. (2005) filters out those ratings that are not in the majority amongst other ones by using the Iterated Filtering approach, hence a review is determined as a false review if the estimated rating lies outside q and (1-q) quantile of the underlying rating distribution. However, this model has some drawbacks as it is only effective when the majority of ratings are fair. Further, these kind of mechanisms are only applicable to scale ratings and not useful for written (text) reviews.

Another mechanism popular among several rating sites is the use of a system based on special fraud detection algorithms. However, those system are mainly focused on detecting so
called ‘manipulative’ false reviews, published by businesses themselves, either to improve their reputation or damage the reputation of a competitor. They work by logging IP addresses and placing cookies. Still, a good written negative review part of a blackmailing threat probably goes unnoticed by such the system. The German consumer foundation Stiftung Warentest states, according their examination, only two out of 13 false reviews were identified as such by the review sites that were tested, which supports the claim that these mechanisms have only limited success in identifying false reviews.

3. Mitigating effect of false reviews

To mitigate the effect of false reviews the concept of “feedback reputation” can be applied.

Feedback reputation is proposed to measure the credibility of users in terms of providing truthful feedbacks. By means of this method, the overall rating of a Hotel, for example, is calculated as the weighted average of all its ratings, whereas the weight of the overall rating is determined by the feedback reputation of the user who provides the rating. As a consequence, ratings provided by users with low feedback reputation will have less impact on the overall rating score.

To determine the feedback reputation for a user, several factors have to be taken into account. Keates (2007) identifies three characteristics that usually indicate fake reviews, these are ratings that differ greatly from the average, mentioning nearby properties as superior and having written about only one hotel and only used the reputation system to post the one review in question. Likewise, Dellarocas (2003) agrees on the point that the most important factor in considering the credibility of a reviewer is the overall number of reviews posted (positive or negative). Accordingly, the results by O’Connor (2010) suggest the standard deviation for single reviews are significantly higher, what indicates more extreme responses. However, he also argues that this in itself does not prove that single reviews per se have to be false, but that such reviews tend to be more extreme, which in turn may be explained by the well known phenomenon of extremely negative or extremely positive customers are more likely to provide feedback (O’Connor, 2010). The findings from Ott et al. (2012), in turn, seem to support the claim given a clear reduction for deception prevalence in reviews for TripAdvisor after removing reviews written by first-time reviewer. Hence, given these conditions, a feedback reputation for a user can be established on the basis of his overall number of reviews posted and the time the user has been active on the review site.

The conclusions given by the foregoing outline of the three main approaches by reputation systems to detect and handle false reviews are summarized in the following table (table 1), highlighting the advantages and disadvantages of each approach.

In the light of the findings from the preceding analysis of the of reputation system mechanisms the result suggest that although these mechanisms provide some degree of efficiency for dealing with false reviews they are not considered overall effective for protecting against blackmailing threats. Still, from the three approaches discussed the one that appears most useful for preventive matters is the first, namely increasing the cost of reviews. For reactive purposes the third approach, mitigating the effect of false reviews provides a good foundation.

<table>
<thead>
<tr>
<th>Approach</th>
<th>Rating</th>
<th>Review</th>
<th>Advantage</th>
<th>Disadvantage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Approach 1</td>
<td>Increasing the cost of reviews</td>
<td>✓</td>
<td>eliminate the threat of speculative attackers</td>
<td>not all reputation systems favor this approach</td>
</tr>
<tr>
<td></td>
<td></td>
<td>✓</td>
<td>lack of anonymity prevent users from engaging in malicious behavior</td>
<td>competitive disadvantage due to small sampling size</td>
</tr>
<tr>
<td>Approach 2</td>
<td>Detection of false review</td>
<td>✓</td>
<td>useful for scale ratings</td>
<td>limited filtering approach is only useful if the majority of ratings are fair</td>
</tr>
<tr>
<td></td>
<td></td>
<td>×</td>
<td>negative filtering proves effective for filtering out suspicious ratings</td>
<td>fraud detection algorithms aimed at detecting manipulative false reviews (not useful for malicious false reviews)</td>
</tr>
<tr>
<td>Approach 3</td>
<td>Mitigating the effect of false reviews</td>
<td>✓</td>
<td>ratings provided by users with low feedback reputation will have less impact on the overall rating score</td>
<td>presumes that users who post more reviews have more credibility</td>
</tr>
<tr>
<td></td>
<td></td>
<td>×</td>
<td>reduces deception prevalence in reviews</td>
<td></td>
</tr>
</tbody>
</table>

Table 1: reputation system approaches

2.3.2 Business Perspective

Given the adverse effects negative reviews can have on business, the problem of blackmailing threats is a widely discussed topic among business owner. As discussed in the previous section, the methods used by reputation sites can not be relied on for providing enough protecting against blackmailing threats. Therefore, the question is how businesses themselves can manage these threats.

As previously pointed out, since blackmailing is a quite recent phenomenon, there is limited scientific research available, none of which were found to provide strategies on this matter. Hence, for the moment the only available strategies concerning the phenomenon of blackmailing in particular, are found in blogs or similar non-scientific literature.

For the matter of this research several media publications by different authors were analyzed in order to elaborate a guiding approach on how to deal with blackmailing threats. The findings from this investigation are summarized in a four stage action model. The four stages are (1) react, (2) document, (3) report and, (4) respond. The different stages are described in the following subsections.

1. React

When the threat is expressed in person it is advised to treat it like a traditional complaint and likewise it is important that the person receiving the threat remains calm and above all professional. Craig (2013) explains that some people are less likely to follow through with their threat once calmer heads prevail, especially if the situation is handled with empathy and professionalism. Thus, trying to offer the customer options, and do everything within reason to find a resolution should be the first approach.

However, there should be no tolerance for sleazy demands. Giving in to these kind of threats can be even more dangerous and have long term repercussions because as Patel (2014) forewarns, blackmailers are frequently connected and share their success and thus, if the word is spread that a business is susceptible to this type of exploitation, it will attract more of this kind of customer.
2. Document

When handling blackmailing threats documentation is key. Patel, hotel owner and founder of smartguests.com, explains that when dealing with guests who come with a negative review threat it is crucial to document everything – every note sent, every call made, everything that was done to conciliate them or to improve the situation (Patel, 2014). It is hard to argue with facts and having the proceeding of the threat clearly documented is also helpful for subsequent actions.

3. Report

Some review systems have introduced mechanisms for business owners to report blackmailing threats. TripAdvisor for example has a protocol for business owners to report threats to its Management Centre before a malicious review is submitted to the site. However, it is very important to submit the report right after the threat has been voiced. TripAdvisor emphasizes that this functionality only works if the malicious review has not been submitted yet. Moreover, they stress that reporting a threat is only effective if the information in the review matches what was included in the blackmail report (TripAdvisor, 2013), this underlines the importance of having a documenting policy.

4. Respond

If despite everything the customer follows through and the negative review appears online there are two options. The preferable case is to have the fake review removed. However, getting a negative review changed or removed on sites such as TripAdvisor can be very hard, or even impossible. The chance to have the review removed is higher when it can be argued that it violates the sites’ terms of use, for example if it uses foul language or can be classified as defamatory, derogatory or similar. Otherwise, removal can only be successful if the business is able to provide conclusive supporting evidence which shows, without a doubt, that the review is factually incorrect. However, as many cases demonstrate the chances of having a false review removed is so low that it is often not even worth the effort. Hence, if removal is not feasible the best option is to write a management response.

Most of the review sites provide a reply facility where businesses can respond to reviews. Responding to reviews, real or fake, positive or negative, should always be part of the review management process (Merritt, 2014). Again when it comes to responding to a fake malicious review, it is critical to do so in a calm and professional manner, respond should be factual and never emotional. Passive-aggressive attitude, sarcastic comments or engaging with the reviewer in a negative way only worsen the situation and in the end reflects poorly on the business itself. Ultimately, responding strategically and carefully is really the key. Jon Hall, founder and CEO of customer review management software GradeUs, advises that in order to neutralize the effect of a fake negative review the best method is to respond to it in a way that makes the situation clearer to third-party (Merritt, 2014). Thus, the response should be addressed at other potential customers and not the attacker. Generally speaking, people value humanity and genuine responses and by responding to reviews potential customers can see that the business is engaged and that it cares about its customers.

According to a study by TripAdvisor, hotels that provide management responses to reviews, both positive and negative, have a 21% increased likelihood of booking inquiries via TripAdvisor, compared to those that do not respond to any. Furthermore, the results suggest that responding to reviews has also a positive effect on the average rating (TripAdvisor, 2014). Consequently, it can be said that with a consistent response policy a business can turn a bad online review into a positive outlook. However, apparently this feature is rarely if even used. In the study conducted by O’Connor (2010) less than 0.5% of the reviews had a management response attached. Craig (2011) talks about a 7% respond rate to negative reviews by hotels. Thus, there is still room for improvement and hoteliers who want to guard themselves from the negative effect of negative reviews should become more proactive in this matter. As emphasized by Looker et al. (2007), while WOM cannot be controlled, it can be managed and must not be ignored.

![Figure 1: Four stage action model](image)

With the presented four-stage-action model (Figure 1) businesses are given a guideline on how they can mitigate the risk of blackmailing threats.

However, in order to have the model work it is crucial to engage in education and training of the staff. As Daniel Edward Craig, founder of online reputation management consulting firm Reknown points out “Review threats can put staff in an extremely difficult position. They don’t want to cave in to unreasonable demands, but they also don’t want to be blamed for a bad review.” Therefore it is essential to have the staff informed about the phenomenon and train them on how to react to such threats (Craig, 2013). He suggests the business should set up clear guidelines so that employees know where they stand in such situations, the options they have and that upper management will support their decisions. Foremost, it has to be assured that in any case these situations are handled with professionalism.

In order to further decrease the risk of social media blackmailing it is pivotal to engage customers to leave positive feedback, since by empowering the normally silent majority of satisfied customers to leave feedback the threat of a single negative false review can be drastically reduced (Whitfield, 2014).

2.3.3 Social Media Blackmailing: what not to do
In the previous section a four stage action model was presented that can be used for businesses as a guide for what to do when facing blackmailing threats. This section provides some suggestions about what not to do regarding blackmailing threats.

First of all a blackmailing threat should not be ignored, to ignore the threat of a negative review is a poor strategy. As the findings from this study suggest, negative reviews can be critical for a business online reputation and therefore should not be overlooked. Furthermore, ignoring a negative review, false or not, provides to other potential customers the image that the business does not care when customers have negative experiences which adds to a poor reputation.

Business owners faced with multiple negative reviews may be tempted to rebalance their reputation with false positive reviews, for example by paying people to post positive reviews. However, the research on reputation systems revealed that sites such as TripAdvisor or Yelp have sophisticated mechanisms in place to filter out such false reviews. Thus, when a business is caught engaging in such activity the consequences can be severe. Therefore, businesses should not try to game the system.

On the other hand, they should also not try to combat the review system. Driven by the fear for the damage that could be done by a single unrepresentative review on a big community some businesses have started to implement clauses in their customer contracts to prevent customers from writing anything negative about the contractor (Carter, 2014). However, many professionals advert that taking such measures does not favour a business reputation, in contrary as Vels Jensen, chief marketing officer at review site Trustpilot points out, such dishonest behaviour rather turns away potential customers, who value transparency and their right to voice (Carter, 2014).

3. CASE ANALYSIS
In order to give the problem of social media blackmailing a practical view and to compare it with above discussed theoretical findings a case study is provided drawn from two examples from the hotel industry. The aim is to analyze the practical approaches on how the hotels deal with blackmailing threats which is then compared with the theoretical findings from the literature review. The cases are derived by means of interviews, which consisted of six open questions (see Appendix A).

Case 1: Radisson Blue Hotel, Lucerne
The four star Radisson Blu Hotel, located at the shore of the lake Lucerne, is a leisure and business hotel offering 189 rooms and is part of the Carlson Rezidor Hotel Group.

The hotel has not experience any extreme kinds of blackmailing attacks so far. However, they confirm that there are guests who are trying to sneak up benefits or perks by complaints on site, sometimes even deliberately looking for something to object upon, but it rarely occurs that a direct threat is expressed without an underlying complaint, such a threat often comes expressed in a positive way, for example something like "We are looking forward to your great house and views of the lake, we read so much good about you ... maybe there will be the possibility of an upgrade ... we would of course mention you on all internet portals with a positive review". The hotel tries to meet the customers demands where possible, but if a guest has already voiced a review 'threat', positive or negative, it is documented. Nevertheless, the majority of the threats normally occur on site and come along with a complaint.

The findings from theory suggested that a customer with a complaint, justified or not, should be treated in a professional and respectful manner, even if it seems obvious that the customer has a malicious motivation. The Radisson Blu hotel ensures this behavior by having Codes of Business ethics every staff member has to follow. In principle, they treat every complaint as justified and take every objection seriously, for this matter the hotel has a specific service concept that build on a “make it right” strategy. This means that every employee has the right to treat complaints at his own discretion in order to satisfy the guest within the possible. However, the codes also include moral guidelines that imply not to give in into extortion threats.

To ensure that every complaint is professionally handled every employee of the Carlson Rezidor Hotel Group receives a training where he is introduced to the company’s Code of Business Ethics and trains the judgment of the employee. There is also specific training on guest complaints. Still, a general rule remains that blackmailing is not accepted. Thus, in the worst case, a malicious false negative review, the hotel director always supports the employees decisions.

In general the hotel supports the principle of freedom of speech reputation systems are build on and therefore does not hold review sites responsible for possible negative reviews. Furthermore, the hotel makes use of the reply function provided by review sites and responds to negative as well as positive feedback from their clients.

Case 2: Iberostar Playa de Muro, Majorca
The IBEROSTAR Playa de Muro is a beachfront family hotel part of the IBEROSTAR Hotels & Resorts Group.

With its 473 rooms it is more than twice as big as the Radisson Blu hotel. Therefore, it comes to no surprise that they are much more regularly faced with nitpicking customer who try to take advantage of any flaw to get hold of a discount, upgrade or freebie by threatening with a bad review. Staff reports around twenty such cases per week. Again, the scheme seem to be that the customer rather tries to exploit an existing dissatisfaction to extort a certain benefit as a compensation. The hotel has no clear guidelines to follow when facing such a demand. When
the front staff is confronted with such a situation, the general rule is to first try to solve the problem in a professional and calm manner, show the client what he has booked and paid for and discuss possible options. However, the front staff is not entitled to make decisions that involve discounts or freebies. If the customer insists on such a demand, or if the guest is behaving in a very aggressive manner, he is forwarded to the head of public relations, or in very extreme cases to the hotel director. The hotel director has a rather intimidating approach to solve such cases. He would invite the guest to a private conversation, offering the guest to take seat while he himself remains in a standing position. He would then ask the customer to describe the situation while listening calmly. If the complaint is perceived as justified, he would offer the possible options to solve the issue in a reasonable manner. If the client acts ignorant and insists on an unreasonable demand he will be released without a unification. Every incident is documented by the staff in a handwritten “notification book”, later these comments are passed on to the public relations department who introduces the information into the software system. Furthermore, whenever an unjustified complaint is expressed with a threat to post a negative review, the hotel reports the case to TripAdvisor. In terms of online reputation management the hotel has both a proactive and reactive approach. The staff of the hotel has recently undertaken a training on opinion management, especially on how to engage satisfied customers to write positive reviews, as discussed in theory this can be very important in order to reduce the risk of any negative review. Moreover, any negative feedback is always addressed with a manager response. This intersection of the real world practice drawn from the two cases and the suggestions from previously discussed theory is presented in table 1.

4. CONCLUSION

Due to the developing of social media and its interconnected environment that has led to growing consumer empowerment, businesses face severe challenges when it comes to their online reputation. As it has been highlighted by this research, online customer ratings are gaining increasing importance and relevance in the decision-making process of the consumer which leads to a growing concern for the repercussion negative reviews can have on a businesses reputation, especially in service industries such as the hospitality industry. In the past years a growing number of cases where reported where customers, misusing their power, try to take advantage of the vulnerability of hotels’ online reputation, extorting them by threatening with a bad review, unless demands like refunds and upgrades are fulfilled. The purpose of this review was to gather and assess relevant information on this specific issue of social media blackmailing in order to gain a deeper understanding of the phenomenon and provide a systematic approach for businesses to deal with the problem.

An approach was made to try to identify the type of customer engaging in such attacks as well as the underlying motivation. The findings from section 2.2.1 suggest that customers who engage in social media blackmailing are likely to be ‘extortionists’ or ‘opportunistic complainers’ motivated foremost by personal gains.

The second aim of the paper was to provide information on how businesses can deal with blackmailing threats. The findings from this research suggest that although a blackmailing threat can be critical, businesses are far from helpless.

<table>
<thead>
<tr>
<th>Theory</th>
<th>Radisson Blu Lucerne</th>
<th>Iberostar Playa de Muro</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Stage 1: React</strong></td>
<td>✓ General guidelines on how to treat guests also applicable for complaining customers (Code of Business Ethics)</td>
<td>✓ Intention to find a solution for the customers’ satisfaction</td>
</tr>
<tr>
<td>remain calm and professional, offer options, try to find a solution</td>
<td>✓ Moral guidelines to not give in into extortion</td>
<td>✓ No specific guidelines</td>
</tr>
<tr>
<td>no tolerance for sleazy demands</td>
<td>✓ General staff is expected to treat customer complaints in a professional and calm manner</td>
<td>✓ Intention to find a solution for the customers’ satisfaction</td>
</tr>
</tbody>
</table>

| **Stage 2: Document** | ✓ Documentation via hotel software | ✓ Use of a “notification book” for notifications about incidents |
| write down every detail from the incident | ✓ | |

| **Stage 3: Report** | ✓ Aware of the report blackmailing option of review sites | ✓ Reporting unjustified complaints and threats to TripAdvisor |
| the incident to the relevant review sites | ✓ | |

| **Stage 4: Respond** | ✓ Responding policy for both positive and negative reviews | ✓ Responding policy for negative reviews |
| Leave a management response to the review | ✓ | |

| **Education/training** | ✓ General training on the Companies Code of Business Ethics and "make it right strategies" | ✓ Opinion management training for encouraging positive feedback |
| ✓ Specific training on customer complaints | ✓ | |

Table 2: overview of findings from case studies

It was found that, in the context of the hotel sector, there is a possibility to mitigate the risk of malicious false reviews if following certain principles. The main principles were summarized and laid out in the four-stage-action model which can be used as a guideline when confronted with a blackmailing threat. When comparing the proposed model with practices from theory, the model proved itself as compatible, yet both of the studied hotels applied practices similar to the model in their daily operation.

Concluding, this research highlights the importance of taking the threat of blackmailing seriously, but also showed that it is a threat that can be managed if hotels are proactive and apply practices to mitigate the risk of false negative reviews.

While negative reviews, especially when badly managed, can be highly damaging, Vermeulen and Seegers (2009) put forward that a single negative online review generally does not cause much harm, whereas a single positive online review can do a lot of good. Therefore, a general takeaway from this study is the importance for business to engage in encouraging positive feedback from their satisfied customers and manage the negative feedback by providing a management response, showing that the business cares about its customers.
5. LIMITATIONS AND RECOMMENDATION FOR FURTHER RESEARCH

It is a common problem that research on new social phenomena face a shortage of academic literature while there is an abundance of non-academic literature. This was also the case in this study of the phenomenon of social media blackmailing. However, whenever non-academic references have been used in this paper, the selection of sources was based as much as possible on the reputation and validity of the source. Accordingly, several of the ideas and the model presented in this study are subject for further study and empirical analysis. Foreground, an empirical testing of the elaborated four-step-action model is necessary in order to validate its usefulness. Also the proposed theory of the type of customer engaging in blackmailing threats and the underlying motivations require empirical testing.

Another limitation is given by the sample of the case study, both hotels are four-star category and part of a bigger hotel chain, thus no generalization can be made. Therefore research with a broader sample is suggested.

Finally, the normative style of the second part of this paper is also due to the mentioned lack of scientific reference and is highlights the need to establish some ground for further research.

6. ACKNOWLEDGMENTS

Many thanks to Dr. E. Constantinides for giving me the opportunity to observe such an interesting topic and providing me useful feedback. I am also thankful for the collaboration Maike Steggemann and her support as a friend. Finally, but first in my heart, I am deeply grateful for my parents for their continued moral and financial support throughout my studies.

7. REFERENCES


The Telegraph (2014, May 8). *Hotels ‘hostage to TripAdvisor blackmailers’*. The Telegraph


Appendix A Interview questions

1. How often do experience cases of blackmailing threats?
2. Is there a specific scheme customers proceed when voicing a blackmailing threat?
3. How do you proceed with such incidents? Do you have specific guidelines?
4. Do you train your staff to deal with such incidents?
5. Do you have any preventive mechanisms to avoid the occurrence of such incidents in advance?
6. Do you believe review sites such as TripAdvisor or HolidayCheck should take more action against blackmailing incidents?